

Faculté des bioingénieurs

**Développement d'une méthode de traitement
des séries temporelles des satellites Sentinel-1
permettant la discrimination des écosystèmes
forestiers du bassin du Congo**

Auteur :	Julien Herinckx
Promoteurs :	Pr. Pierre Defourny Dr. Julien Radoux
Lecteurs :	Pr. Quentin Ponette Vincent Istace
Année académique :	2018-2019
Bioingénieur :	Gestion des forêts et des espaces naturels

Remerciements

Je souhaite par la présente note, remercier les nombreuses personnes qui ont contribué de près ou de loin à la bonne réalisation de ce mémoire de fin d'études, que ce soit vis-à-vis de la collecte de données sur le terrain, du traitement de ces données ainsi qu'à la rédaction de ce mémoire.

Je tiens premièrement à remercier mes promoteurs de mémoire, Pr. Pierre Defourny et Dr. Julien Radoux, de m'avoir offert cette opportunité de pouvoir travailler sur un sujet captivant incluant une expérience sur le terrain hors du commun. J'aimerais également les remercier pour leur expertise autant théorique que pratique m'ayant guidé vers la bonne réalisation de ce mémoire.

Je voudrais, en second lieu, remercier Vincent Istace et Pr. Quentin Ponette d'avoir accepté d'être lecteurs de ce mémoire.

Je souhaite ensuite remercier chaleureusement Vincent, Arnaud, Alain, Laurent, Walter, Denis, Patrick, Dimitri et tous les autres membres de la CIB qui ont tout fait en sorte pour que mon expérience en leur compagnie soit la plus enrichissante et agréable possible. J'ai pu, grâce à eux, faire tant de belles découvertes, m'immerger dans une nouvelle culture, un monde dont je ne connaissais rien.

Je tiens également à remercier Romain et Macaire avec qui je travaillais régulièrement sur le terrain et grâce à qui j'ai pu découvrir la forêt tropicale, dans toute sa splendeur sauvage, ainsi que dans ses détails les plus discrets et intimes.

Je désire ensuite exprimer mes remerciements à tous les membres du laboratoire de géomatique pour leur aide et leurs conseils quotidiens.

J'aimerais remercier ma famille et surtout ma mère, Laurence, pour ses nombreuses corrections et commentaires aussi utiles que nécessaire.

Je souhaite finalement remercier Caroline, Céline, Alexis, Baptiste et Dorian avec qui j'ai passé de nombreuses heures dans notre très cher local C.389. Je les remercie pour leur aide ainsi que leur soutien moral qui nous aura permis de progresser ensemble et de clôturer ce travail ensemble. Je tiens plus particulièrement à remercier Céline pour les nombreux fruits qu'elle nous a offerts, des petits moments de plaisirs quotidiens.

Table des matières

Chapitre 1 : Etat de l'art.....	3
1.1 Description des forêts et écosystèmes forestiers tropicaux	3
1.2 Etat des forêts en République du Congo	5
1.3 Radar à synthèse d'ouverture (SAR)	8
1.4 Les apports de la télédétection satellitaire dans la cartographie forestière.....	10
1.5 Segmentation	14
1.6 Analyse texturale	17
1.7 Classification	19
1.8 Conclusion de la revue de la littérature.....	21
1.9 Objectifs	22
Chapitre 2 : Zone d'étude	23
2.1 Climat	23
2.2 Topographie	25
2.3 Pédologie.....	26
Chapitre 3 : Matériel et méthode.....	27
3.1 Matériel.....	27
3.2 Méthode.....	34
Chapitre 4 : Résultats et discussion	56
4.1 Campagne de terrain	56
4.2 Classification	62
4.3 Proposition d'une nouvelle méthodologie	77
Chapitre 5 : Conclusion	78
Chapitre 6 : Perspective	79
6.1 Indice radar de végétation (RVI)	79
6.2 Angle d'incidence	79
6.3 Interférométrie et mesure de cohérence	80

Liste des acronymes

1A1	Méthode « One Against One »
1AA	Méthode « One Against All »
ALOS	Advanced Land Observing Satellite
BRDF	Bidirectional Reflectance Distribution Function
CCRS	Canada Centre for Remote Sensing
CIB	Congolaise Industrielle des Bois
CNIAF	Centre national d’inventaire et d’aménagement des ressources forestières et fauniques
COMIFAC	Commission des forêts d’Afrique centrale
DEM	Modèle d’élévation de terrain
DOS	Dark object subtraction
ESA	Agence Spatiale Européenne
ESP	Estimation du paramètre d’échelle
EW	Extra Wide swath mode
FSC	Forest Stewardship Council
GLCM	Matrice de co-occurrence des niveaux de gris
IW	Interferometric Wide swath mode
JAXA	Japan Aerospace Exploration Agency
JAROS	Japan Resources Observation System Organization
JPL	Jet Propulsion Laboratory
LIDAR	Laser Detection and Ranging
LV	Variance Locale
MRF	Markov Random Field
MPM	Méthode de maximisation des marginaux postérieurs
MRS	Segmentation multirésolution
NASA	National Aeronautics Space Administration
NESZ	Noise Equivalent Sigma Zero
NGA	National Geospatial-Intelligence Agency
OBIA	Object-Based Image Analysis
OFAC	Observatoire des forêts d’Afrique centrale
PALSAR	Phased Array L-band Synthetic Aperture Radar
PFBC	Partenariat pour les Forêts du Bassin du Congo
RBF	Fonction Radiale de Base
RCM	RADARSAT Constellation Mission
RCS	Radar Cross section
REDD	Réduction des Emissions provenant de la Dégradation forestière et de la Déforestation
RF	Random Forest
RVI	Radar Vegetation Index
S1A	Sentinel-1A
S1B	Sentinel-1B
SA	Simulated annealing
SAR	Radar à synthèse d’ouverture
SIG	Système d’Information Géographique
SLC	Single look complexe
SLIC	Simple Linear Iterative Clustering

Liste des acronymes

SM	StripMap mode
SNAP	Sentinel Application Platform
SNR	Rapport signal sur bruit
SRTM	Shuttle Radar Topography Mission
SVM	Support Vector Machine
TOPSAR	Terrain observation with Progressive Scan SAR
UFA	Unité forestière d'aménagement
WV	Wave mode

Liste des figures

Figure 1 : Production moyennée annuelle en grume provenant du rapport de l'Observatoire des Forêts d'Afrique Centrale (OFAC, 2017).....	7
Figure 2 : Résolution temporelle des satellites Sentinel-1. Source de l'image : https://sentinel.esa.int/documents/247904/3339704/Sentinel-1-Revisit-Coverage-Frequency-Geometry-2018.jpg	8
Figure 3 : Méthode d'acquisition des données par sous-fauchée selon le mode IW. Figure provenant du site internet de l'ESA : https://sentinel.esa.int/web/sentinel/user-guides/sentinel-1-sar/acquisition-modes/interferometric-wide-swath	9
Figure 4: Polarisation circulaire dans le sens horloger (R). Figure créée par Dave3457 issue de la page https://en.wikipedia.org/wiki/Circular_polarization	11
Figure 5: Polarisation circulaire dans le sens anti-horloger (L) Figure créée par Dave3457 issue de la page https://en.wikipedia.org/wiki/Circular_polarization	11
Figure 6 : Interprétation visuelle des différents éléments de la biomasse aérienne forestière permettant la rétrodiffusion des ondes électromagnétiques (Le Toan et al., 2002)	12
Figure 7 : Schématisation du principe de segmentation basée sur les variables aléatoires de Markov Random Field	15
Figure 8 : Classification linéaire en utilisant un SVM (Malik and Malhotra, 2016)	20
Figure 9 : Classification non linéaire en utilisant un SVM (Malik and Malhotra, 2016)	20
Figure 10 : Classification des climats de Köppen-Geiger révisée par Beck et al., 2018.	24
Figure 11: Diagramme climatique de Pokola (https://fr.climate-data.org/afrique/congo-brazzaville/sangha/pokola-52835/)	24
Figure 12 : La marque rouge de l'image correspond à l'emprise des données capturées par le satellite Sentinel-1A sur l'orbite absolue 20029. Echelle : 1 :6 000 000. Image de fond SIO, NOAA, U.S. Navy, NGA, GEBCO (Google Earth).....	27
Figure 13 : Interprétation visuelle des quatre principales strates rencontrées sur le terrain. Les mesures de hauteur [mètres] ne sont présentes qu'à titre indicatif et font référence aux cas les plus fréquemment rencontrés.	31
Figure 14 : Les zones non exploitées, lorsqu'elles ne sont soumises à aucune autre perturbation naturelle, présentent généralement une structure climacique, la canopée y est homogène et fermée. Certaines formations forestières peuvent, au contraire des structures mono-climaciques, présenter une structure méta-climacique. Les stades climaciques se succèdent induisant une canopée pouvant être hétérogène.	32
Figure 15 : Les zones exploitées présentent de fortes variations de hauteur sur de faibles distances induisant une texture particulière à l'image. Les essences en présence sont différentes de celles des structures climaciques et sont majoritairement pionnières à post-pionnières, selon la date de la perturbation.....	32

Liste des figures

<i>Figure 16 : Analyse de la hauteur maximale des arbres depuis un transect de 270 mètres en forêt secondaire à Triplochiton scleroxylon (Ayous) située à la position 16°24'7"E – 1°33'51"N</i>	<i>33</i>
<i>Figure 17 : Analyse de la hauteur maximale des arbres depuis un transect de 430 mètres en forêt primaire dominée par le Gilbertiodendron dewevreii (Limballi) située à la position 16°27'52,8"E – 1°38'32,8"N.....</i>	<i>33</i>
<i>Figure 18 : Évolution du NDVI (vert) et de la probabilité d'acquisition d'une image composite sans nuages (barres grises) sur l'année 2015 - Pokola, Congo. Estimation à partir des données de réflectance acquises par le capteur MERIS (ESA/CCI-LC, UCLouvain, 2015)</i>	<i>34</i>
<i>Figure 19 : Comparaison de trois images issues de la différence des coefficients de rétrodiffusion [dB] des deux images S1 moyennées sur les périodes de maximum et de minimum de production de biomasse végétale. Les images résultant de cette différence sont ventilées selon un écart-type de ± 1 (a), $\pm 1,5$ (b) et ± 2 (c).</i>	<i>35</i>
<i>Figure 20 : Ces deux figures représentent l'espace de recherche des deux algorithmes k-moyennes (à gauche) et SLIC (à droite). Dans la méthode conventionnelle k-moyennes, les distances sont calculées depuis chacun des groupements sur tous les pixels de l'image. L'algorithme SLIC limite son espace de calcul des distances du centre de chacun des groupements jusqu'aux pixels compris dans une surface de $2S * 2S$. Il faut noter que la taille présumée d'un superpixel est de $S * S$, indiquée par le carré pointillé (Achanta et al., 2012).....</i>	<i>36</i>
<i>Figure 21 : Comparaison de 3 segmentations réalisées à partir de l'algorithme SLIC selon un facteur de fusion de 25 % et de compacité de 100. Le nombre d'itérations est fixé à 10 pour ne modifier qu'une seule variable, le facteur de forme. Celui-ci est modifié entre les différentes segmentations comme suit : (a) Facteur de 10, (b) Facteur de 15 et (c) Facteur de 20. Les images du dessus correspondent à l'image originale segmentée par superpixels, celle du dessous aux données spectrales moyennées par superpixels. Composition colorée : VV (Rouge), VH (Vert), VV/VH (Bleu)</i>	<i>37</i>
<i>Figure 22 : Comparaison de 3 segmentations réalisées à partir de l'algorithme SLICO selon un facteur de forme de 7 et de fusion de 10 %. Seul le nombre d'itérations est modifié entre les segmentations : (a) 5 itérations, (b) 10 itérations, (c) 50 itérations. Les images du dessus correspondent à l'image originale segmentée par superpixels, celle du dessous aux données spectrales moyennées par superpixels. Composition colorée : VV (Rouge), VH (Vert), VV/VH (Bleu).</i>	<i>38</i>
<i>Figure 23 : Comparaison de 3 segmentations réalisées à partir de technique de segmentation multirésolution selon un paramètre d'échelle de 30, un indice de compacité de 1. Seul le paramètre de forme est modifié entre les segmentations : (a) paramètre de forme de 0,2, (b) paramètre de forme de 0,4, (c) paramètre de forme de 0,6. Les images du dessus correspondent aux frontières des objets, délimitées par la segmentation, celles du dessous, aux données spectrales moyennées par objet. Composition colorée : VV (Rouge), VH (Vert), VV (Bleu).....</i>	<i>39</i>
<i>Figure 24 (au-dessus) : Évolution de la variance locale en fonction du paramètre d'échelle selon la méthode développée par Drăguț, Tiede, et Levick (2010) et améliorée par Drăguț et al. (2014). Cette évaluation de LV est réalisée sur base de l'ensemble de la segmentation de la zone d'étude.....</i>	<i>42</i>
<i>Figure 25 (à gauche) : Résultat de la segmentation par ESP2. Les zones anthropisées aux alentours de Pokola (Rouge) ont été extraites de la segmentation à partir d'une classification supervisée sur la valeur moyenne du coefficient de rétrodiffusion de la bande VV, VH et VV/VH ainsi que sur les paramètres texturaux du contraste, de la corrélation et des moyennes GLCM. Echelle : 1 : 75 000.</i>	<i>42</i>
<i>Figure 26 : Moyenne des écarts-types de la hauteur maximale de la canopée distribuées selon les dates d'exploitations des parcelles échantillonnées</i>	<i>43</i>

Liste des figures

<i>Figure 27 : Matrice de corrélation des différentes caractéristiques texturales de l'image.</i>	45
<i>Figure 28 : Organigramme des traitements effectués sur base des données SAR-C.</i>	47
<i>Figure 29 : Schéma suivi pour la réalisation de la segmentation stratifiée.</i>	50
<i>Figure 30 : Stratification des différentes classes résultant de la segmentation. Les classes situées au sommet de la figure se superposent aux classes situées en bas de la figure.</i>	50
<i>Figure 31 : Paramètres des segmentations fines des forêts sèches et marécageuses</i>	50
<i>Figure 32 : Visualisation de l'impact du passage d'un filtre à majorité sur une image en entrée (a) et son résultat après application de l'outil (b) par analyse des 4 pixels avoisinants le pixel d'intérêt. Figure issue du panneau d'aide des fonctions de ArcMap10.3.</i>	52
<i>Figure 33 : Visualisation de l'impact de l'utilisation de l'outil "nettoyage des limites" sur une image en entrée (a) et son résultat après application de l'outil (b). Figure issue du panneau d'aide des fonctions de ArcMap10.3. ..</i>	52
<i>Figure 34 : Visualisation des forêts mixtes anciennement exploitées sur base des données SAR-C [composition VV, VH, VV/VH] (en haut à droite), de terrains (en haut à gauche) et de transect (en bas).</i>	56
<i>Figure 35 : Visualisation des forêts mixtes récemment exploitées sur base des données SAR-C [composition VV, VH, VV/VH] (en haut à droite), de terrain (en haut à gauche) et de transect (en bas).</i>	57
<i>Figure 36 : Visualisation des parasoleraies sur base des données SAR-C [composition VV, VH, VV/VH] (en haut à droite), des données terrain (en haut à gauche) et des transects (en bas).</i>	57
<i>Figure 37: Visualisation des marécages à Bayas sur base de données SAR-C [composition VV, VH, VV/VH] (en haut à droite, des données terrains (en haut à gauche) et des transects (en bas).</i>	58
<i>Figure 38 : Visualisation des marécages à Raphia sur base des données SAR-C [composition VV, VH, VV/VH] (en haut à droite), des données terrains (en haut à gauche) et des transects (en bas).</i>	58
<i>Figure 39 : Visualisation des peuplements de bubingas sur base des données SAR-C [composition VV, VH, VV/VH] (en haut à droite), de terrain (en haut à gauche) et de transect (en bas).</i>	59
<i>Figure 40 : Visualisation des eyangas sur base des données PALSAR [composition HH, HV et HH/HV], de terrains (en haut) et de transect (à droite). Echelle 1:25.000.</i>	59
<i>Figure 41 : Identification visuelle des peuplements de limballis à partir des données Sentinel-1 [Composition VV, VH, VV/VH] (au-dessus), Sentinel-2 [bande 6 (rouge), 11 (vert), 8 (bleu) (au milieu) et sur le terrain (en dessous) selon trois milieux différents : Plateaux (a), milieux rivulaires (b) et milieux marécageux (c).</i>	60
<i>Figure 42 : Distribution des valeurs des coefficients de rétrodiffusion (bande 1 : VV, bande 2 : VH, bande 3 : VV/VH) pour trois classes de peuplements de limballis : plateaux : vert, rivulaires : rouge et marécageux : bleu.</i>	61
<i>Figure 43 : Données acquises sur la hauteur maximale de la canopée via un transect créé en forêt à limballi ...</i>	61
<i>Figure 44 : Identification visuelle des prairies inondables sur base des données SAR-C [composition VV, VH, VV/VH] (en haut), le profil issu du transect (en haut à droite), et sur le terrain (à droite)</i>	61

Liste des figures

<i>Figure 45 : Evolution du coefficient de Kappa et de la précision générale selon quatre classifications par pixels</i>	62
<i>Figure 46 : Visualisation des cours d'eau de la Ndoki sur base des données SAR-C [Composition VV, VH, VV/VH], le tracé du cours d'eau est indiqué en surbrillance (à gauche) et des données terrains (au-dessus). Les images drones ont été capturées le 18 juin 2018 par Laurent Teillier, gérant du groupe Africadrones.</i>	71
<i>Figure 47 : Variation du coefficient de rétrodiffusion en fonction de l'angle d'incidence d'acquisition des données SAR. La figure démontre un comportement différent de ce coefficient selon la rugosité de la surface d'intérêt. Pour un angle inférieur à 25°, les surfaces modérément rugueuses présentent un coefficient supérieur aux surfaces rugueuses (NASA,s.d). Source de la figure : https://history.nasa.gov/JPL-93-24/p53b.htm</i>	72
<i>Figure 48 : Visualisation de la distorsion radiométrique présente à l'extrémité de la sous-fauchée IW3 sur une longueur de 20,5 km. Donnée SAR-C [Composition VV, VH, VV/VH].</i>	73
<i>Figure 49 : Schématisation du problème permettant de calculer l'angle d'incidence correspondant aux distorsions radiométriques</i>	73
<i>Figure 50 : Variation du coefficient de rétrodiffusion des polarisations VV, VH et VV/VH en fonction de l'angle d'incidence. Ces données sont approximées à partir d'une polynomiale de troisième degré. La partie de la courbe mise en surbrillance jaune correspond au signal du zone marécageuse tandis que celle en surbrillance rouge correspond au biais radiométrique.</i>	74
<i>Figure 51 : Données d'ombrage (au-dessus) et d'exposition des pentes (en dessous), extraites à partir des données SRTM.</i>	75
<i>Figure 52 : Visualisation de la variation radiométrique due à la superposition des fauchées IW2 et IW3</i>	76
<i>Figure 53 : Biais radiométrique sur la surface équivalente radar (RCS) mesurée en fonction de NESZ (Piantanida, 2017)</i>	76
<i>Figure 54 : Proposition de traitement des données permettant l'élaboration de la classification par objet présentant les meilleures performances</i>	77
<i>Figure 55 : Scénario d'observation des satellites Sentinel-1</i>	80

Liste des cartes

<i>Carte 1 : Trois principaux massifs forestiers de la République du Congo : Kouilou-Mayombe (Sud), Niari-Chillu (Centre) et Nord Congo (Nord) (Mertens et al., 2007).</i>	5
<i>Carte 2 : Topographie des concessions forestières de la CIB.</i>	25
<i>Carte 3 : Données pédologiques des sols des concessions forestières de la CIB. Données pédologiques mises à disposition par la cellule d'aménagement de la CIB.</i>	26
<i>Carte 4 : Zones échantillonnées lors de la campagne de terrain. Au total 52 échantillons furent collectés, à raison d'un échantillon par zone (symbolisée par l'étoile verte).</i>	30
<i>Carte 5 : Distribution des mesures de rugosité de la canopée pour chacune des zones visitées. Les dates mentionnées en arrière-plan de la figure correspondent aux années d'exploitation des parcelles.</i>	44
<i>Carte 6 : Classification par objet de l'occupation du sol des concessions forestières de l'UFA Pokola et Kabo. ...</i>	66
<i>Carte 7 : Zoom de la classification sur les zones anthropisées. La segmentation réalisée sur base de l'algorithme MRS-ESP2 est indiquée par les traits grisés.</i>	67
<i>Carte 8 : Zoom sur les zones anthropisées. La segmentation réalisée sur base de l'algorithme MRS-ESP2 est indiquée par les traits grisés</i>	67
<i>Carte 9 : Zoom de la classification sur les zones marécageuses et les prairies inondables. La segmentation fine des marécages, réalisée sur base de l'algorithme MRS, est indiquée par les traits grisés.</i>	68
<i>Carte 10 : Zoom sur les zones marécageuses et les prairies inondables. La segmentation fine des marécages, réalisée sur base de l'algorithme MRS est indiquée par les traits verts</i>	68
<i>Carte 11 : Zoom sur la classification des forêts sèches et marécageuses. La segmentation, basée sur l'algorithme MRS est indiquée par les traits grisés.</i>	69
<i>Carte 12 : Zoom sur les forêts sèches et marécageuses. Les deux segmentations ont été réalisées sur base de l'algorithme MRS. Les traits rouges correspondent à la segmentation des marécages et les traits verts à celle de la forêt.</i>	69
<i>Carte 13 : Zoom sur la classification des eyangas.</i>	70
<i>Carte 14 : Zoom sur les eyangas sur base des données SAR-L [Composition colorée : HH, HV, HH/HV]</i>	70

Liste des tableaux

<i>Tableau 1 : Principaux éléments de la biomasse aérienne forestière permettant la rétrodiffusion des ondes électromagnétiques (Le Toan et al., 2002).....</i>	<i>12</i>
<i>Tableau 2 : Principales caractéristiques texturales décrites par Haralick (1973).</i>	<i>17</i>
<i>Tableau 3 : Angle d'incidence des données acquises selon le mode IW. La troisième sous fauchée (IW3) correspond aux données utilisées pour la zone d'étude.....</i>	<i>28</i>
<i>Tableau 4 : Ensemble des images satellites ayant permis l'identification des zones d'intérêt lors de la campagne de terrain. Les données S2 ont uniquement été utilisées dans le cadre de cette campagne afin de faciliter la différenciation des écosystèmes forestiers, la composition [RedEdge, SWIR2, NIR] permettant d'apporter un maximum de contraste entre les différentes formations forestières.</i>	<i>29</i>
<i>Tableau 5 : Comparaison du coefficient de Braun-Blanquet avec le pourcentage de recouvrement d'une essence d'intérêt par projection de la surface de son houppier au sol.</i>	<i>32</i>
<i>Tableau 6 : Synthèse des performances respectives de la segmentation multirésolution, SLIC et SLICO.</i>	<i>40</i>
<i>Tableau 7 : Paramètres conditionnant l'algorithme ESP.....</i>	<i>41</i>
<i>Tableau 8 : Résultat de comparaison des moyennes de deux jeux de données, l'un fixé sur les forêts anciennement exploitées, l'autre fixée sur les forêts récemment exploitées, évalués selon le 95^{ème} et le 99^{ème} percentile.</i>	<i>44</i>
<i>Tableau 9 : Jeux de données utilisées pour les différents processus de classification.</i>	<i>51</i>
<i>Tableau 10 : Distribution des vrais/faux positifs/négatifs selon les accords et désaccords des valeurs réelles et prédites par la classification.....</i>	<i>53</i>
<i>Tableau 11 : Degré d'accord et valeur de Kappa proposé par Landis and Koch (1977)</i>	<i>54</i>
<i>Tableau 12 : Ensemble des données utilisées en intrants pour réaliser les différentes classifications par pixels. Les performances respectives de chacune des classifications et classes sont également reprises. Le nombre d'aires d'entraînement de chacune des classes est indiqué entre parenthèses à côté de la classe correspondante.....</i>	<i>62</i>
<i>Tableau 13 : Matrice de confusion de la classification mêlant 6 couches objet [VV, VH, VV/VH, Dissimilarité, Moyenne GLCM et Variance GLCM] et le modèle d'élévation de terrain [SRTM].</i>	<i>64</i>

Introduction

Les forêts humides tropicales et subtropicales couvrent une surface de 19.800.000 km², soit 13,5 % de la surface terrestre émergée. Comprise dans des latitudes inférieures à 30° nord et sud de l'équateur, ces forêts abritent l'une des biodiversités les plus riches sur terre (Bergonzini et Lanly, 2000). Outre la richesse biologique, les forêts tropicales sont responsables de nombreux services écosystémiques appartenant à différentes classes telles que le prélèvement, la régulation ou la protection. La première concerne la collecte de bois et de produits forestiers non ligneux. La seconde aborde les cycles de régulation du carbone (fixation et stockage) et de l'eau. La troisième, finalement, est liée à la protection des milieux. En effet, la présence d'un couvert végétal permet d'empêcher les coulées de boues, de diminuer l'intensité des inondations et de limiter les phénomènes d'érosion (Chevassus-Au-Louis et Pirard, 2011).

Bien que l'état des connaissances sur les forêts tropicales se développe de plus en plus, ce dernier reste toutefois lacunaire. Cette connaissance lacunaire peut s'expliquer par plusieurs raisons. Tout d'abord, la collecte de données en forêts tropicales est particulièrement ardue due à leur surface importante et leur accessibilité limitée. De plus, leur constitution complexe et leur dynamique naturelle apportent une certaine difficulté dans le traitement de ces informations (van der Sanden et Hoekman, 1999).

Outre cette difficulté d'acquisition des données, les écosystèmes forestiers tropicaux sont sujets à une augmentation des pressions liée notamment aux pratiques de déforestation ainsi qu'aux phénomènes de changements climatiques. Afin de répondre à ces pressions croissantes, les écosystèmes forestiers tropicaux sont de plus en plus fréquemment étudiés. Cette multiplication des études s'explique également par l'augmentation du nombre d'acteurs impliqués directement ou indirectement dans les modifications et transformations de ces écosystèmes. Suite à l'augmentation récente des contraintes et défis liés aux écosystèmes forestiers, il est nécessaire de caractériser ces derniers avec une précision accrue.

L'amélioration des connaissances sur les forêts tropicales aura de multiples impacts potentiels. Dans le meilleur des cas, celle-ci permettra de répondre à l'ensemble des pressions dont est sujet la forêt tropicale. Une meilleure compréhension des mécanismes écologiques et biologiques de ces milieux fournirait des outils facilitant l'implémentation des programmes REDD+ et des projets de conservation de la biodiversité. En outre, elle pourrait faciliter la mise en accord de politiques liées aux changements climatiques. D'un point de vue plus pratique, une telle compréhension permettra également de cibler les actions les plus importantes à effectuer afin d'arriver à une gestion plus durable des forêts tropicales.

Le récent lancement des satellites européens Sentinel-1 (A : 3 avril 2014, B : 25 avril 2016) et Sentinel-2 (A : 23 juin 2015, B : 7 mars 2017) offre une excellente opportunité pour travailler sur ces milieux de vaste étendue. En effet, la télédétection satellitaire est un outil particulièrement performant ayant fait ses preuves par le passé pour caractériser et identifier les différents écosystèmes forestiers (Chambers et al., 2007), évaluer les modifications du couvert forestier (Verhegghen et al., 2016) ou pour étudier l'évolution de la déforestation (Blanc et al, 2016). L'apport nouveau des satellites Sentinel-1 et Sentinel-2 permettra la collecte de données synoptiques, systématiques et avec une grande répétition temporelle.

L'étude se déroulera dans le bassin du Congo, le bassin versant du fleuve Congo. Cette zone abrite la seconde plus grande forêt tropicale au monde qui couvre une surface de 2 millions

Introduction

de kilomètres carrés (Bwangoy et al., 2010) et comprend 10 pays : l'Angola, le Burundi, le Cameroun, le Gabon, la République centrafricaine, la République du Congo, la République démocratique du Congo, le Rwanda, la Tanzanie et la Zambie.

Chapitre 1 : Etat de l'art

Ce chapitre a pour vocation de dresser l'état des connaissances sur les matériels et méthodes ayant permis la description des écosystèmes forestiers tropicaux du bassin du Congo. Ce chapitre comptera quatre sections. La première de ces sections veillera à réaliser une description générale des différents écosystèmes forestiers du bassin du Congo et se clôturera par un sommaire historique de la gestion des forêts de la République du Congo. La seconde partie décrira les modes d'acquisitions et autres caractéristiques des différents capteurs dont les données ont été utilisées dans le cadre de cette étude. La troisième section de ce chapitre décrira les différentes méthodes d'analyse et de traitement des images radars utilisée pour la discrimination des écosystèmes forestiers. La quatrième et dernière section de ce chapitre reprendra les différents objectifs de cette étude.

1.1 Description des forêts et écosystèmes forestiers tropicaux

L'Observatoire des forêts d'Afrique centrale (OFAC) publie en janvier 2016 un rapport ayant pour fonction la construction d'une typologie forestière harmonisée en Afrique centrale. La création d'un tel rapport a permis le développement d'une classification régionale des forêts du bassin du Congo.

Le terme « forêt » est soumis à plusieurs définitions variant selon les pays. Afin de pouvoir représenter de manière non biaisée l'ensemble des forêts du monde, il est nécessaire de trouver une définition qui fasse consensus. Cependant, compte tenu de la variabilité des écosystèmes forestiers, il est difficile de trouver une telle définition qui satisfasse la majorité.

Néanmoins, la FAO a émis une définition du terme « forêt » qu'elle décrit comme suit :

Le terme forêt correspond à « *des terres occupant une superficie de plus de 0,5 hectare avec des arbres atteignant une hauteur supérieure à cinq mètres et un couvert arboré de plus de 10 %, ou avec des arbres capables d'atteindre ces seuils in situ* ».

Cette définition est régulièrement utilisée comme référence pour la classification forestière. L'OFAC appuie cette définition en distinguant les forêts des terrains occupés par des arbres hors forêt. Cette distinction, basée sur la densité du couvert forestier, se précise comme suit pour l'Afrique centrale :

Sont catégorisés comme forêts les terrains forestiers s'ils :

- *ont un couvert arboré feuillu d'une hauteur minimale de 5 m*
- *ont une densité de couvert arboré supérieure ou égale à 25 %*
- *sont installés naturellement ou semi-naturellement.*

Finalement, l'OFAC propose une typologie des forêts permettant de distinguer quatre formations forestières. Cette typologie se base sur plusieurs critères liés à la fermeture du couvert, à l'hydromorphie, à la phénologie ou encore à l'état de conservation.

Forêts denses humides sempervirentes

Les formations forestières présentent une strate arborée majoritairement composée d'arbres feuillus sempervirents. Elles se rencontrent dans des zones de forte pluviométrie (2000-3000 mm/an) le long de la bordure atlantique du Gabon et Cameroun et à l'Est du bassin du Congo (de Namur, 1990). Le couvert arboré reste donc verdoyant toute l'année grâce au renouvellement continu du feuillage. La strate arborée supérieure est uniforme, dense et

dominée par une ou quelques essences d'ombre à régénération abondante et bien représentée dans toutes les strates (SPIAF, 2007). La strate arbustive, limitée dans son développement, est souvent recouverte d'épiphytes. La strate herbacée y est toujours absente. Les caractéristiques indispensables pour la catégorisation de ces forêts sont :

- *un couvert arboré sempervirent supérieur à 60 % (bien distribué en âge) composé d'essences sempervirentes*
- *une hauteur d'arbres élevée (40-50m)*
- *un couvert arbustif ($1 < x \leq 40$ %)*
- *la présence d'épiphytes.*

Ces formations forestières peuvent être distinguées selon les espèces sempervirentes présentes dans l'étage arboré (ex : *Gilbertiodendron dewevrei*, *Julbernardia seretii*, *Brachystegia laurentii*,...).

Forêts marécageuses

Les formations forestières marécageuses sont des formations évoluant sur un sol inondé en permanence ou, durant une grande partie de l'année, par de l'eau douce. Ces formations présentent également une strate arbustive. Elles se distinguent ainsi d'autres types de couverts forestiers par :

- *un couvert arboré de feuillus sempervirents, dense, fermé (>60 %) de hauteur comprise entre 30 et 40 m*
- *un sol inondé (> 4mois) par de l'eau douce*
- *un couvert arbustif ouvert à dense (15-40 %).*

Ces forêts peuvent être distinguées selon les associations spécifiques. Un certain type de forêt marécageuse, typique en Afrique Centrale, peut être ici distingué : les *Raphiales*. Ces formations forestières marécageuses se trouvent le long de cours d'eau ou de plans d'eau fermés et sont caractérisées par la dominance de *Raphia sp.*.

Forêts périodiquement inondées

Ces formations forestières d'Afrique Centrale se développent sur un sol périodiquement inondé (1 à 4 mois) d'eau douce. Elles sont constituées d'une strate arborée dense de feuillus (30-40m) semi-décidus et présentent également une strate arbustive ouverte à fermée :

- *un couvert arboré dense, fermé (>60 %)*
- *un sol périodiquement inondé (1-4 mois) par de l'eau douce*
- *un couvert arbustif ouvert à dense (15-40 %).*

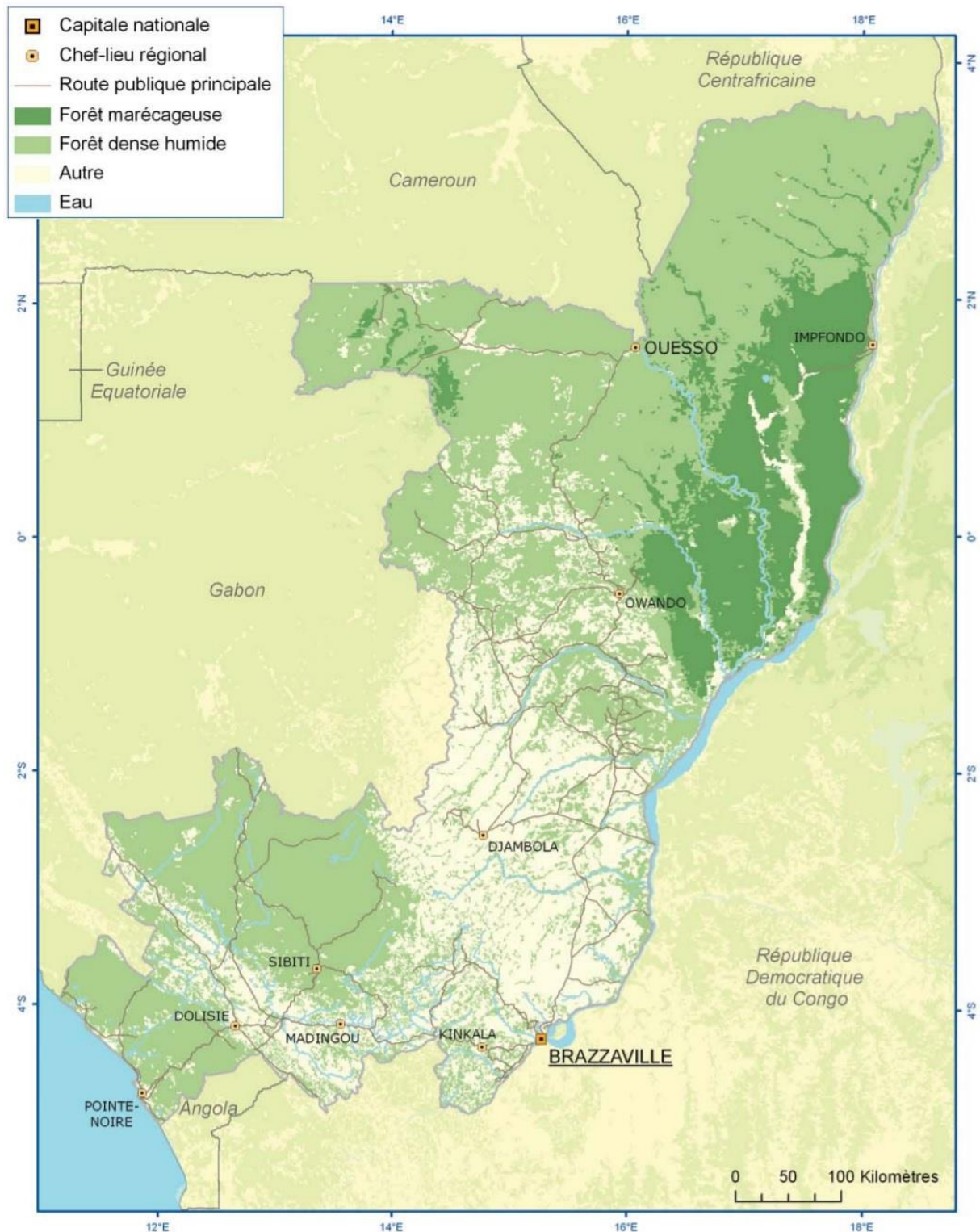
Forêts dégradées

Les forêts de terre ferme du bassin du Congo sont soumises à de nombreux facteurs de dégradation, essentiellement liés à l'accroissement démographique dans un contexte de pauvreté et de faible développement socio-économique. Ces forêts subissent alors une réduction de leur capacité à fournir des biens et services (FAO, 2011).

D'un point de vue physiologique, les forêts dégradées ne sont pas aussi développées et présentent une richesse spécifique moindre par rapport aux forêts intactes. De plus, les espèces arborées qui dominent ces formations ne sont pas les mêmes que celles que l'on retrouve au sein des formations climaciques. Ces dernières se matérialisant principalement par la présence de parasoliers et autres essences pionnières.

1.2 Etat des forêts en République du Congo

Les forêts congolaises s'étendent sur une surface de plus de 20 millions d'hectares ventilés selon 3 massifs distincts : Kouilou — Mayombe, Niari — Chillu et Nord Congo. Une telle couverture forestière correspond à un potentiel ligneux d'environ 560 millions de m³ de volume exploitable (Onkagui, s. d.).



Carte 1 : Trois principaux massifs forestiers de la République du Congo : Kouilou-Mayombe (Sud), Niari-Chillu (Centre) et Nord Congo (Nord) (Mertens et al., 2007).

Au cours des dernières décennies, les programmes de gestion des forêts congolaises ont bien évolué :

La République du Congo proclame son indépendance le 15 août 1960. Les premières lois nationales sont publiées. Ces dernières concernent notamment la gestion forestière et traitent essentiellement de la récolte et de la commercialisation du bois ainsi que de la perception des taxes sur cette ressource. L'aménagement forestier n'y est que peu abordé (Onkagui, 2003).

Les premiers plans d'aménagement forestiers sont proposés dans les années 1970. Les thématiques de gestion durable sont difficiles à intégrer, l'état ayant pour principal but la rentabilité de la production de bois. De plus, celui-ci possède le monopole sur les domaines touchant la commercialisation du bois, la sylviculture et la constitution des forêts. (Onkagui, 2003)

Depuis, la cessation du monopole mené par l'état permet l'élaboration de nouveaux articles dans le code forestier promouvant une gestion plus durable des écosystèmes forestiers. (Onkagui, 2003)

Les derniers chiffres en date indiquent que la République du Congo présente une couverture forestière de 23,5 millions d'hectares (soit 69 % du territoire). Le pays est peu sujet aux phénomènes de dégradation et déforestation, car seul 0,05 % sont déforestés par an. Cela représente une surface d'environ 12. 000 ha/an (CNIAF, 2015). Les causes de la déforestation sont principalement dues aux quatre facteurs indiqués ci-dessous (REDD, 2016) :

- les pratiques non durables d'agriculture itinérante sur brûlis,
- la production et la consommation non durable de bois énergie,
- l'exploitation forestière non durable et/ou illégale et,
- le développement urbain.

Ces différents facteurs sont aggravés par d'autres facteurs sous-jacents tels que le manque de vision commune d'aménagement du territoire, l'augmentation démographique, la pauvreté rurale et l'absence d'alternatives aux sources d'énergie adaptées aux populations à faible revenu (REDD, 2016).

La République du Congo a cherché à s'adapter à ces problèmes de déforestation en améliorant la gestion forestière via la mise en place de concessions forestières et d'aires protégées. Le territoire forestier congolais sujet aux exploitations est actuellement découpé en une cinquantaine de concessions forestières. À ce jour, trente concessions forestières, soit 76,04 % de la surface forestière concernée par l'exploitation, sont soumises à un aménagement. De ces trente concessions, quinze d'entre elles sont déjà soumises à un plan d'aménagement. Afin d'avoir une production plus durable de ces forêts, quatre concessions forestières sont certifiées par le label *Forest Stewardship Council* (FSC) et couvrent 2.478.943 ha, ce qui positionne la République du Congo parmi les premières places des pays tropicaux avec les plus grandes superficies de forêts tropicales humides certifiées FSC au monde (REDD, 2016).

Un réseau d'aires protégées d'une surface de 3,6 millions d'hectares (10,6 % du territoire national) a également été développé. Ces aires ont pour objectif d'assurer la conservation du système forestier, pour promouvoir la recherche scientifique ainsi que le tourisme de vision (COMIFAC/OFAC, 2014). En 2006, l'OFAC indique que le secteur forestier contribue à 5,6 % de la valeur du PIB de la République du Congo. En 2007, ce sont plus de 7.400 emplois qui ont été

créés directement en lien avec la gestion forestière. Comme l'indique la Figure 1, la production en bois n'a cessé d'augmenter au cours des dernières décennies.

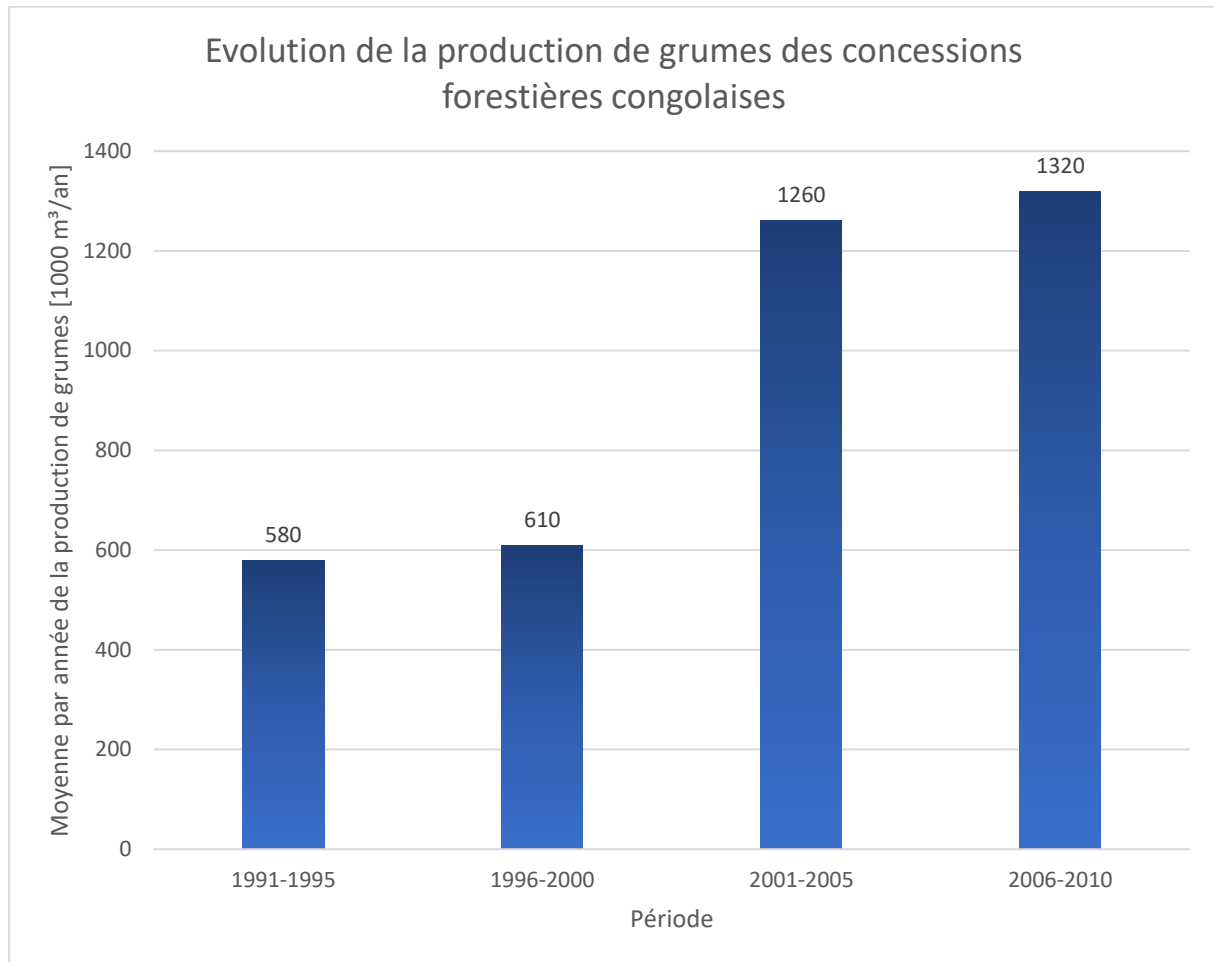


Figure 1 : Production moyennée annuelle en grume provenant du rapport de l'Observatoire des Forêts d'Afrique Centrale (OFAC, 2017).

Compte tenu de ces différents facteurs, les ressources forestières représentent de grands enjeux, autant au niveau national que mondial. Afin de mieux gérer ces ressources, il est nécessaire de caractériser ces dernières avec grande précision.

Les techniques de télédétection permettent de répondre à la principale contrainte d'acquisition de donnée en forêt tropicale : l'accessibilité. Elles se définissent comme suit : «la télédétection est l'ensemble des connaissances et des techniques permettant de déterminer certaines caractéristiques physiques et biologiques de points observés à partir de mesures effectuées à distance, sans contact matériel avec ceux-ci (CNTRL, 2012).» Ces techniques font régulièrement usage de satellites ou d'autres véhicules aéroportés munis de capteur afin d'acquérir des données sur une surface ou un phénomène d'intérêt. On distingue deux grands types de capteurs, l'un passif permettant l'acquisition d'images optiques, l'autre actif permettant l'acquisition de données radar ; ce dernier s'inscrivant directement dans la dynamique de recherche de ce mémoire.

1.3 Radar à synthèse d'ouverture (SAR)

La télédétection radar permet l'acquisition de données de manière active (en opposition avec l'acquisition passive de données qui s'appuie sur la luminance de la surface d'intérêt). L'onde émise par la technologie radar est une onde électromagnétique dont la fréquence varie selon la bande utilisée. Dans le cadre de cette étude, les bandes principales utilisées seront la bande C (4-8 GHz) et la bande L (1-2 GHz). L'image résultant de l'acquisition des données est donc une représentation de la rétrodiffusion par la surface d'intérêt pour une longueur d'onde particulière dans le spectre électromagnétique (Agence spatiale européenne, 2018).

Ce mémoire se concentrera essentiellement sur l'utilisation d'images radar générées par les satellites Sentinel-1A (S1A) et Sentinel-1B (S1B). Ces derniers, situés sur une orbite héliosynchrone à 693 km d'altitude, fournissent des données gratuites à raison d'une image tous les 3 jours pour l'Europe, 6 jours pour le reste du monde (Agence spatiale européenne, 2018). D'autres données, acquises par le satellite japonais ALOS-2, seront également utilisées. Cette plateforme contient l'instrument PALSAR-2 permettant d'acquérir les données en bande L. Le satellite ALOS-2 fut lancé le 24 mai 2014. Il est placé sur orbite héliosynchrone à une altitude variant entre 636 km et 639 km et effectue son cycle en 14 jours (ESA, 2019).

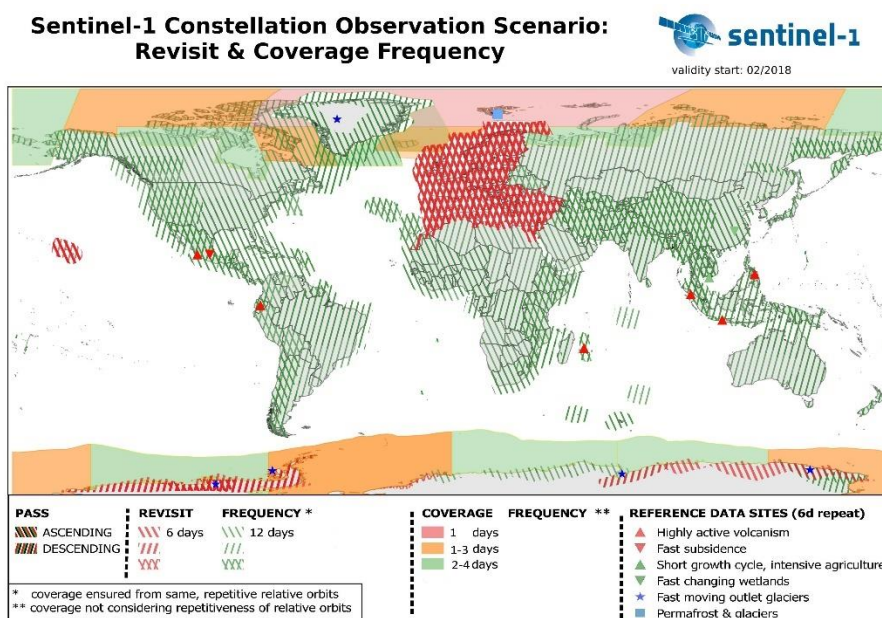


Figure 2 : Résolution temporelle des satellites Sentinel-1. Source de l'image : <https://sentinel.esa.int/documents/247904/3339704/Sentinel-1-Revisit-Coverage-Frequency-Geometry-2018.jpg>

1.3.1 SAR-C

Le principal instrument présent sur les satellites Sentinel-1 est le radar à synthèse d'ouverture en bande C (SAR-C). Le capteur de ces satellites permet l'acquisition de données à une fréquence de 5405 MHz selon différentes polarisations, celles-ci pouvant être simples (HH ou VV) ou doubles (HH+HV ou VV+VH). La première lettre de cette symbolique des polarisations indique que le satellite émet une onde polarisée horizontalement (H) ou verticalement (V). La seconde quant à elle, indique que le satellite capte une réponse de la surface d'intérêt selon une polarisation horizontale (H) ou verticale (V). La polarisation double est dès lors capable de coupler une émission d'onde horizontale avec une acquisition d'onde verticale et inversement.

Outre les différentes polarisations, les instruments SAR-C sont capables d'acquérir des données selon quatre modes différents (Agence spatiale européenne, 2018) :

- **StripMap mode (SM)** : L'acquisition des données se fait selon une fauchée de 80 km pour une résolution spatiale approximant les 5 mètres. L'antenne du satellite est dirigée de manière à acquérir des données selon un angle azimutal fixe et un angle off-nadir presque fixe. La surface d'intérêt est illuminée par un flux de pulsation constante.
- **Interferometric wide swath mode (IW)** : Cette méthode est la plus couramment utilisée pour l'acquisition de données sur les surfaces terrestres. Le mode IW réalise trois sous-fauchées utilisant la technique *Terrain observation with Progressive Scan SAR* (TOPSAR). Cette méthode capture les données par pulsation en changeant le faisceau de l'antenne de manière cyclique entre les trois sous-bandes. Une image est produite pour chaque sous-fauchée et pour chaque polarisation pour un total de 6 images par produit IW. La méthode IW couplée avec la technique d'acquisition TOPSAR permet d'obtenir des données sur des fauchées de grande amplitude (250 km) tout en permettant d'avoir des images à performance radiométrique améliorée. La technique TOPSAR permet d'avoir des images dont le ratio $\frac{\text{Signal}}{\text{Bruit}}$ (SNR) est uniforme sur toute la surface d'acquisition.
- **Extra wide swath mode (EW)** : Cette méthode d'acquisition est principalement utilisée sur les zones couvertes par la glace, les pôles et les banquises. La méthode d'acquisition se base également sur la technique TOPSAR, mais en réalisant 5 sous-fauchées. La fauchée principale fait 410 km de large et permet l'acquisition de données à une résolution spatiale de 20 mètres sur 40 mètres.
- **Wave mode (WV)** : Cette méthode consiste en l'acquisition de données sur de multiples vignettes de 20 km sur 20 km selon un mode de polarisation simple (HH ou VV) avec une résolution spatiale de 5 mètres. Chacune de ces vignettes est stockée séparément dans une image indépendante. L'acquisition des données sur ces vignettes est réalisée tous les 100 km en alternant deux angles d'incidence (approximativement 23° et 36°). En conséquence, les vignettes définies par un même angle d'incidence sont séparées d'une longueur de 200 km.

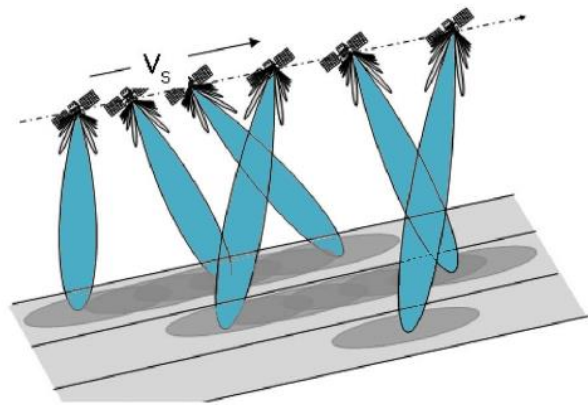


Figure 3 : Méthode d'acquisition des données par sous-fauchée selon le mode IW. Figure provenant du site internet de l'ESA : <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-1-sar/acquisition-modes/interferometric-wide-swath>

1.3.2 PALSAR

Le radar PALSAR-2 est un des instruments du satellite japonais *Advanced Land Observing Satellite* (ALOS) également appelé *Daichi*. Les objectifs d'ALOS sont multiples. Bien qu'il se concentre principalement sur la cartographie du Japon et des pays avoisinants, il est également utilisé pour réaliser une observation continue des ressources naturelles mondiales (JAXA, 1997).

L'instrument PALSAR-2 est un capteur utilisant la fréquence de la bande-L (1270 MHz). Il présente une fauchée comprise entre 250 et 350 km variant selon le nombre de capteurs et influant sur la résolution spatiale. Les données mises à disposition sont des mosaïques d'images moyennées par année selon les polarisations HH et HV et à une résolution de 25 mètres. La mission PALSAR fut développée de manière conjointe entre la *Japan Aerospace Exploration Agency (JAXA)* et la *Japan Resources Observation System Organization (JAROS)*.

1.4 Les apports de la télédétection satellitaire dans la cartographie forestière

Compte tenu du lancement récent des satellites Sentinel-1, les études permettant la caractérisation des forêts et autres milieux naturels à l'aide des données SAR-C sont relativement peu nombreuses.

Cependant, les données radars ont, par le passé, permis la mise en évidence de nombreux éléments, qu'ils soient liés au couvert forestier, aux régimes hydriques ou à la topographie.

1.4.1 Ecosystèmes forestiers

La cartographie forestière peut se décliner selon différents axes. La description des écosystèmes forestiers est l'un d'entre eux et fut réalisée par le passé pour différents pays appartenant notamment à l'Afrique subsaharienne (Gond et al. 2015 ; Mayaux et al., 1999). Afin de réaliser des classifications correctes, les principales sources de données utilisées proviennent d'images optiques (MODIS, Sentinel-2, Landsat 8). La classification résultant du traitement de ces images est ensuite affinée à partir d'images radar (PALSAR, SAR-C) ou LiDAR (GLAS). Contrairement aux images radar, la qualité des images optiques peut grandement être affectée par des perturbations atmosphériques. Il existe de nombreuses méthodes permettant de corriger ces perturbations atmosphériques. Ceux-ci peuvent, par exemple, être limités en compilant des images d'une seule et même zone d'étude à des dates différentes. Ainsi, les parties de l'image d'une date donnée qui sont polluées par des nuages et autres perturbations atmosphériques sont corrigées à partir d'images d'autres dates.

La *dark object subtraction* (DOS) est une seconde méthode permettant de limiter les effets de l'atmosphère. Elle consiste en l'analyse de la réflectance des pixels des objets sombres tels que les étendues d'eau profonde. Ces pixels présentent un niveau de réflectance égal ou très proche de zéro. Ce signal, presque nul, est considéré comme étant dû aux effets de la diffusion atmosphérique. En conséquence, la soustraction de la valeur de réflectance de ces pixels à l'ensemble des pixels de l'image permettrait de limiter les biais radiométriques dus à l'atmosphère (Gilmore et al., 2015).

Certaines études menées sur des zones de grande dimension nécessitent l'utilisation de plusieurs images. Afin de réduire les variations intrinsèques à l'acquisition de chacune des images, ces dernières peuvent être normalisées par ajustements de l'anisotropie *cross-track* via la fonction de distribution de la réflectivité bidirectionnelle (BRDF). Cette fonction permet de quantifier la distribution géométrique de la radiance résultant d'une source de lumière incidente provenant d'une direction quelconque (Shell, 2004). L'utilisation de ces méthodes permet de réduire les variations radiométriques inter- et intra-tuiles qui sont associées aux variations atmosphériques, phénologiques ainsi qu'aux angles solaires (Gond et al., 2015).

Bien que les méthodes de correction des effets atmosphériques soient nombreuses, l'acquisition de données de qualité sur les écosystèmes tropicaux reste un défi, certaines

zones étant couvertes par un écran nuageux de manière quasi permanente. La technologie radar permet de répondre à cette absence de données due aux effets atmosphériques.

Bien que l'utilisation d'image radar pour améliorer les classifications réalisées à partir d'images Sentinel-2 soit une méthode couramment utilisée pour la cartographie forestière, d'autres études ont privilégié l'utilisation seule des données radar pour classer les forêts (Alena et al., 2016; Grover et al., 1999; Oliver, 2000; van der Sanden and Hoekman, 1999).

Ces études ont permis de mettre en évidence les paramètres et données les plus significatifs pour la réalisation d'une classification correcte. van der Sanden and Hoekman, (1999) réalisèrent la comparaison de deux jeux de données, le premier provenant de la *Canada Centre for Remote Sensing* (CCRS SAR), le second de la *National Aeronautics and Space Administration* (NASA) en collaboration avec le *Jet Propulsion Laboratory* (JPL). De ces deux jeux de données, les bandes X et C furent analysées selon les polarisations HH et VH, VV et HV pour CCRS SAR, tandis que les bandes C, L et P furent analysées selon les polarisations HH, VH, VV, RR, LR, LL pour NASA/JPL.

RR, LR et LL sont des polarisations circulaires ; R indique une polarisation dont l'onde progresse dans le sens horloger (Figure 4), observé depuis l'émetteur tandis que L indique une polarisation dont l'onde progresse dans le sens anti-horloger (Figure 5) (Berthier, 2012).

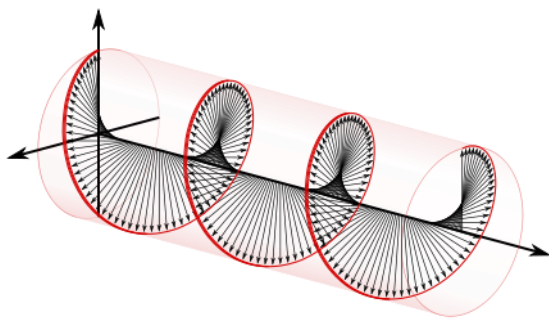


Figure 4: Polarisation circulaire dans le sens horloger (R).
Figure créée par Dave3457 issue de la page https://en.wikipedia.org/wiki/Circular_polarization

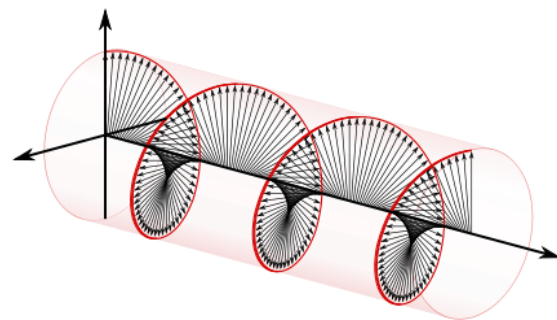


Figure 5: Polarisation circulaire dans le sens anti-horloger (L).
Figure créée par Dave3457 issue de la page https://en.wikipedia.org/wiki/Circular_polarization

Dans le cas de vols aéroportés, la comparaison directe des résultats issus de différents jeux de données est difficile. Outre, les différences au niveau des fréquences et polarisations, cette difficulté s'explique également par les différentes lignes de vols que suivent les capteurs. En effet, les données provenant des polarisations HH et HV ne vont pas être reliées à l'exacte même zone que celles des données VH et VV. De plus, ces régions d'intérêts, observées à des angles d'incidence différents, présenteront des données texturales différentes (van der Sanden et Hoekman, 1999).

Cette comparaison de différents jeux de données a cependant permis d'identifier que les combinaisons des bandes L et P se sont montrées plus efficaces à la classification des forêts exploitées, des forêts secondaires et des zones non forestières que toutes les autres combinaisons des bandes C. De ces combinaisons, la combinaison C-LL, P-VH et P-TP présentait le meilleur potentiel de classification (van der Sanden et Hoekman, 1999). Cette difficulté de classification des forêts par la bande C pourrait s'expliquer par la fréquence de son onde.

Outre la fréquence, l'étagement de la végétation d'un milieu influe sur le comportement des ondes électromagnétiques émises par le satellite. Ces dernières, selon leur fréquence, seront réfléchies par différents éléments de la canopée. Selon le type de la formation végétale, la densité du feuillage ainsi que celle des branches sera différente et impactera la manière avec laquelle l'onde réagit face à ces surfaces. Le Tableau 1 indique que les ondes électromagnétiques de bande C sont réfléchies par les feuilles et branches de faibles dimensions. Ces ondes ne pénètrent donc que faiblement la canopée. Les ondes de la bande L, quant à elles, pénètrent davantage la canopée et ne sont arrêtées que par les éléments biologiques de taille plus importante tels que les branches.

Tableau 1 : Principaux éléments de la biomasse aérienne forestière permettant la rétrodiffusion des ondes électromagnétiques (Le Toan et al., 2002).

Bande	X	C	L	P	VHF
Principaux rétrodiffuseurs en forêt	Feuilles, brindilles	Feuilles petites branches	Branches	Branches et tronc	Tronc



Austrian pine

X band
 $\lambda = 3 \text{ cm}$



L band
 $\lambda = 27 \text{ cm}$



P band
 $\lambda = 70 \text{ cm}$



VHF
 $\lambda > 3 \text{ m}$



Figure 6 : Interprétation visuelle des différents éléments de la biomasse aérienne forestière permettant la rétrodiffusion des ondes électromagnétiques (Le Toan et al., 2002)

1.4.2 Zone humide

Les forêts tropicales sont régulièrement sillonnées par des zones marécageuses et humides présentant des phénomènes hydrologiques particuliers. Ces milieux étant intimement liés aux forêts tropicales, il est nécessaire de pouvoir les caractériser avec autant de précision que les forêts sèches. Ces zones humides sont définies en tant que « surface de terre servant de transition entre les systèmes terrestres et aquatiques où la nappe d'eau est généralement proche de la surface du sol ou que le sol est recouvert par une nappe d'eau superficielle » (Cowardin et al., 1979). Une étude menée par Bwangoy et al. (2010) utilisa plusieurs jeux de données optiques (Landsat) et radar (JERS-1 SAR) ainsi que des indices de topographie pour permettre la cartographie de ces milieux. Bwangoy favorisa l'utilisation des bandes optiques 4 (0,75-0,90 μm), 5 (1,55 -1,75 μm) et 7 (2,09-2,35 μm), ces dernières étant moins sensibles aux effets atmosphériques que les bandes à plus courte longueur d'onde (Ouaidrari et Vermote, 1999). Ces bandes permettent de mettre en évidence les milieux saturés en eau ainsi que les flux transpiratoires de la végétation, deux informations pertinentes pour l'identification des zones humides (Bwangoy et al., 2010).

Afin de capturer au mieux la variation du régime hydrique, les données radar (bande-L) furent collectées pendant deux périodes correspondant au minimum et maximum du fleuve Congo (Bwangoy et al., 2010). Le volume total de rayonnement réfléchi peut être modélisé en utilisant l'équation du transfert radiatif du soleil (1990). Cette modélisation traite la canopée, les troncs et la surface du sol comme des couches qui vont rétrodiffuser et atténuer l'énergie micro-onde incidente (Alsdorf, et al., 2001 ; Melack, 2004). Cette équation se définit comme suit :

$$\sigma_t^\circ = \sigma_c^\circ + \sigma_s^\circ + \sigma_m^\circ + \sigma_d^\circ$$

Avec :

- σ_t° Volume total rétrodiffusé
- σ_c° Volume rétrodiffusé par la canopée
- σ_s° Volume rétrodiffusé par la surface du sol
- σ_m° Volume rétrodiffusé par les multiples interactions entre la canopée et le sol
- σ_d° Volume rétrodiffusé par les interactions entre les troncs et le sol

$\sigma_t^\circ, \sigma_c^\circ, \sigma_s^\circ, \sigma_m^\circ$ et σ_d° sont exprimés en $\frac{m^2}{sr \cdot m^3}$

L'utilisation de la bande-L selon un angle d'incidence de 37° et une double polarisation horizontale permet que les composantes σ_c° et σ_d° contribuent de manière similaire au volume total réfléchi (Wang et al., 1995).

Le rayonnement frappant la surface d'une zone humide typique est réfléchi par la présence de la nappe d'eau superficielle et subit une rétrodiffusion selon un phénomène de double rebond entre troncs et sols, ce qui crée un contraste entre les régions inondées et non inondées permettant la caractérisation des zones humides (Mayaux et al., 2002).

Les effets de la topographie, distribués en 9 facteurs dérivés d'un modèle digital d'élévation (DEM), permettent de mettre en évidence les zones humides qui sont considérées comme des éléments linéaires correspondant à l'alignement de fond de vallée. Ces facteurs sont liés à la pente, l'ombre portée par le relief, l'orientation des pentes de la zone d'intérêt, sa convexité de profil et de plan, sa courbure longitudinale, sa courbure en coupe transversale et sa courbure maximale et minimale (Wood, 1996). Ces facteurs permettent de décrire les variations des systèmes hydriques ainsi que la probabilité qu'un flux d'eau provenant d'une pluie incidente se dirige vers une direction particulière. L'utilisation de ces facteurs topographiques indique dès lors la probabilité qu'une zone particulière soit considérée comme zone humide (MacMillan et al., 2000).

Le couplage des données radar, optique et topographique permet, via un algorithme de classification, d'estimer la probabilité par pixel de présence ou non d'une zone humide (Breiman et al., 1984).

Une fois les différents milieux d'intérêt identifiés, ces derniers peuvent être classifiés selon plusieurs paramètres. Les classifications peuvent se réaliser sur la simple valeur de chacun des pixels ou sur base d'objets, délimités au préalable sur l'image. Dans le cas d'une classification par objet, la création desdits objets nécessite de passer par une opération de segmentation de l'image.

1.5 Segmentation

L'homme a pour faculté de plus facilement interpréter les choses qui l'entourent lorsque celles-ci sont découpées en entités discrètes (Goodchild, 2007). L'analyse d'image n'est pas indépendante d'un tel phénomène. C'est ainsi que naquirent les analyses d'images par objet qui permirent d'apporter une nouvelle perspective d'interprétation des images, les objets portant parfois en eux plus d'informations qu'un simple pixel (Fisher, 1997 ; Smith et al., 2007).

L'étape la plus critique permettant de réaliser des analyses d'images par objet est la création desdits objets par l'agrégation de pixels inclus dans le processus de segmentation. La segmentation est définie comme le processus de division des images en régions discrètes ou en objets qui sont homogènes, en considération de leurs caractéristiques spectrales et spatiales (Ryherd et Woodcock, 1996 ; Drăguț et al., 2010).

Plusieurs méthodes sont utilisées pour réaliser les segmentations d'images SAR. Ces méthodes peuvent être divisées en deux classes distinctes :

La première méthode consiste à réduire le bruit dû au speckle directement via le processus de segmentation. Cette méthode a pour avantage d'utiliser les données d'intensité et la structure inhérente des pixels des données SAR. Le speckle est un phénomène inhérent aux images radar et qui leur confère un aspect très bruité à l'apparence poivre et sel. Le speckle dégrade la qualité de l'image et complique les interprétations de cette dernière. Il est causé par des interférences aléatoires constructives et destructives des ondes électromagnétiques déphasées captées par le satellite (Goodman, 2007).

La seconde consiste à appliquer des filtres permettant de limiter le bruit dû au speckle (Choi et Baraniuk, 2001 ; Fosgate et al., 1997). Ce filtre est appliqué sur les données brutes des images SAR et est suivi d'un processus de segmentation similaire à celui des images optiques. Cette méthode a pour avantage d'être rapide, mais présente deux défauts majeurs. Les images SAR présentant un speckle important nécessitent un filtre tout aussi puissant afin de minimiser l'effet dû au speckle. L'utilisation d'un tel filtre a généralement pour impact négatif de détériorer la qualité des segmentations. De plus, il est difficile d'avoir un contrôle de qualité sur le type de filtre à utiliser pour obtenir des segmentations très performantes. Finalement, il arrive fréquemment que l'utilisation d'un filtre performant introduise des relations et calculs complexes pouvant rendre inefficace l'ensemble du processus de segmentation (Gao et al., 2008).

Afin de traiter les problèmes que pose cette dernière méthode de segmentation, de nombreuses études et recherches liées à l'imagerie SAR se sont focalisées sur l'utilisation d'une méthode de segmentation performante incluant la réduction du speckle dans son processus (Lombardo et Macri Pellizzeri, 2003 ; Cao et al., 2005). Lombardo et Oliver (2002) proposent ainsi une méthode exploitant trois approches distinctes.

La première consistait en l'utilisation du maximum de vraisemblance sur le comportement des variables statistiques des données SAR pour segmenter l'image en zones présentant des caractéristiques homogènes.

La seconde approche utilise un test de fusion et division se basant sur les données structurelles de l'image. Un tel test permet la diminution du bruit dû au speckle tout en réalisant une

segmentation de qualité. Ces deux méthodes présentent cependant le défaut d'utiliser l'algorithme de *Simulated Annealing* (SA) pour maximiser la fonction permettant d'estimer la matrice de variance/covariance, le tout nécessitant un grand temps de calcul.

La troisième, finalement, se base sur l'utilisation des valeurs propres de la décomposition des paramètres de polarimétrie. Cette méthode combine les propriétés statistiques, spatiales et structurales tout en limitant le temps de calcul.

Le modèle Markov Random Field (MRF) est l'un des outils les plus fréquemment utilisés pour réaliser des segmentations prenant en compte les relations spatiales entre pixels (Gao et al., 2008). Différents algorithmes peuvent être utilisés avec MRF tels que le modèle à conditions itératives (ICM) (Besag, 1986), la méthode de maximisation des marginaux postérieurs (MPM) (Cao et al., 2005) ou encore l'approche SA (Cook et al., 1996).

Ces trois méthodes présentaient des résultats prometteurs tout en étant entravées par de nombreuses erreurs d'optimisation. Ces erreurs proviennent du modèle MRF lui-même, ce dernier introduisant des paramètres difficilement estimables (Gao et al., 2008).

Malgré ses performances modérées pour la segmentation d'image radar, le modèle MRF est couramment utilisé en combinaison avec des données optiques. Les images optiques ont pour avantage de présenter un ratio signal/bruit plus élevé que les images radar. Cependant, ces données sont souvent altérées par des perturbations atmosphériques. Xu et al., (2011) proposent une méthode de segmentation basée sur la Figure 7.

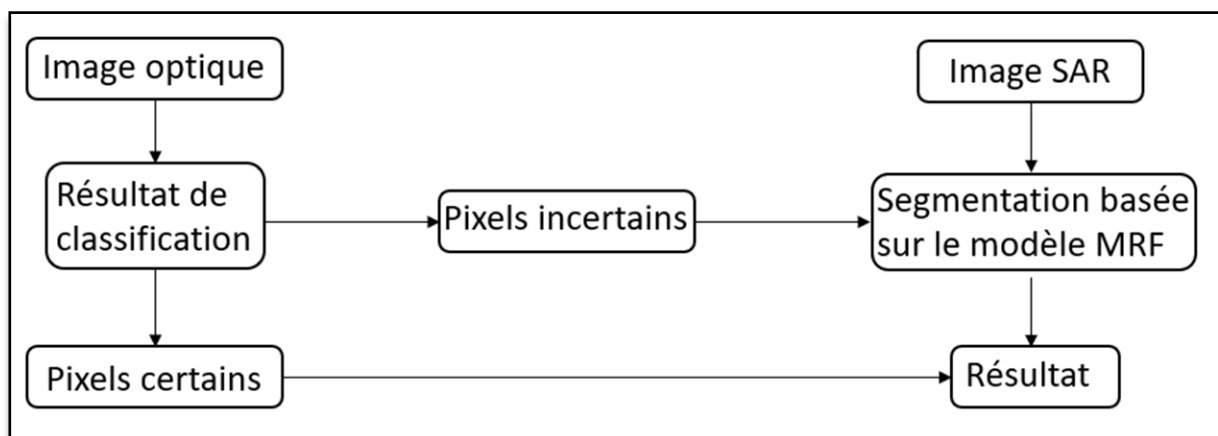


Figure 7 : Schématisation du principe de segmentation basée sur les variables aléatoires de Markov Random Field

Ce modèle se base sur l'utilisation du maximum a posteriori (MAP), de la fonction d'énergie et de la formule de Bayes pour parvenir à la bonne réalisation de la segmentation. La segmentation est donc réalisée, dans ce cas-ci, à partir de données optiques. Les données radar servent à combler les valeurs manquantes des pixels dont le signal est altéré par la présence d'un nuage, d'une ombre ou d'un quelconque effet atmosphérique. Le couplage des données optiques et radar avec le modèle MRF permet de traiter de manière efficiente les informations texturales, tonales, et ce, malgré la présence de speckle. Son principal défaut reste cependant le temps de calcul nécessaire à la production d'un résultat de qualité – 9,459 secondes pour une image carrée de 550 pixels de côté (Xu et al., 2011).

Depuis le lancement des satellites S1A et S1B, de nouvelles approches concernant la segmentation ont été analysées. Clauss et al. (2018) proposent une segmentation des données S1 par la génération de super-pixels via l'algorithme de Clustering par Itération Simple et Linéaire (SLIC) implémentée dans le package Python. SLIC a pour avantage d'être rapide et efficace sur des sites ayant des propriétés différentes, le tout fonctionnant avec un nombre réduit de paramètres ne dépendant pas du type de site. Certains paramètres empiriques (nombre de segments, indice de compacité et de forme) doivent toutefois être ajoutés et adaptés selon le site d'intérêt (Clauss et al., 2018).

Un second algorithme, dérivé de SLIC, fut développé pour répondre à la principale limitation de SLIC : le paramètre de compacité des superpixels. Les images occupées par des zones présentant une importante rugosité texturale étaient segmentées par des pixels très irréguliers tandis que les zones plus lisses étaient segmentées par des superpixels réguliers et homogènes. Il est ainsi difficile de choisir la valeur optimale pour le paramètre de rugosité sans que cela n'affecte l'une ou l'autre zone de l'image. L'algorithme SLICO ne permet plus de modifier le paramètre de compacité, mais adapte celui-ci pour chacun des superpixels formés de manière à ce que ceux-ci soient homogènes sur toute l'image, dans les zones à faible et à fort contraste textural (Achanta et al., 2012 ; EPFL, 2018)

Finalement, plusieurs études récentes, utilisant les données des satellites S1, se basent sur un algorithme de segmentation multirésolution (MRS) (Delgado-Aguilar et al. 2017; Lohberger et al. 2018 ; Clerici et al., 2017 ; Liu et al., 2018). La segmentation multirésolution est une technique de fusion de région d'intérêt permettant de réaliser des séparations maximisant la variabilité inter-objet tout en minimisant la variabilité intra-objet. Le paramètre de fusion de ces régions est donc basé sur un critère d'homogénéité résultant de la combinaison d'un critère spectral et d'un critère de forme. Au total, quatre paramètres sont étudiés pour la réalisation de la segmentation : un paramètre d'échelle, le poids attribué à chacune des couches étudiées, le poids attribué à la valeur spectrale et à la forme des associations de pixels ainsi que le poids attribué à la compacité et à la rugosité des régions. Le paramètre d'échelle, utilisé pour déterminer la taille finale des objets, correspond à l'hétérogénéité maximale autorisée pour la création d'un objet (Liu et al., 2018).

Une fois la segmentation réalisée, de nombreuses caractéristiques peuvent être calculées et extraites de chacun des objets. Une des principales caractéristiques pouvant être extraite de ces objets est l'information texturale. Cette utilisation de la variable « texture » permet souvent d'affiner les classifications, là où les données brutes ne suffisent pas.

1.6 Analyse texturale

Les mesures de texture permettent de mettre en évidence les relations spatiales qui existent entre chacun des pixels composant une image. Haralick (1973) décrit ainsi 14 caractéristiques texturales dérivées des matrices de co-occurrence des niveaux de gris d'une image (GLCM). Les matrices GLCM sont des tableaux décrivant la fréquence d'occurrence de différentes combinaisons des valeurs de niveau de gris de plusieurs pixels sur une image. De ces 14 caractéristiques texturales, cinq d'entre elles, liées à la variance calculée sur la somme des pixels adjacents, la variance calculée sur la différence des pixels adjacents, l'entropie sur la somme et la différence des pixels adjacents, les corrélations des entropies et le coefficient du maximum de corrélation, ne sont généralement pas adoptées pour la génération de classification (Hall-Beyer, 2017).

Les 9 caractéristiques texturales couramment utilisées en tant qu'input pour la génération de classification sont divisées en trois classes (Hall-Beyer, 2017) :

Tableau 2 : Principales caractéristiques texturales décrites par Haralick (1973).

Classe	Caractéristique texturale	Equation
Contrastes	Contraste / Inertie	$\sum_{i,j=0}^{N-1} P_{i,j} (i-j)^2$
	Dissimilarité	$\sum_{i,j=0}^{N-1} P_{i,j} i-j $
	Homogénéité	$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1 + (i-j)^2}$
Statistique descriptive	Moyenne GLCM (μ)	$\mu_i = \sum_{i,j=0}^{N-1} i(P_{i,j})$
	Variance GLCM (σ^2)	$\sigma_i^2 = \sum_{i,j=0}^{N-1} P_{i,j} (i-\mu_i)^2$
	Corrélation GLCM	$\sum_{i,j=0}^{N-1} P_{i,j} \left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right]$
Ordre et arrangement	Angular second moment (ASM)	$\sum_{i,j=0}^{N-1} P_{i,j}^2$
	Energie	$\sqrt{ASM}$
	Entropie	$\sum_{i,j=0}^{N-1} P_{i,j} (-\ln P_{i,j})$

Avec pour variables :

- $P_{i,j}$ Valeur de la probabilité d'occurrence pour la cellule (i,j)
- N Nombre de colonnes ou lignes dans la matrice GLCM
- i Numéro de la ligne dans la matrice GLCM
- j Numéro de la colonne dans la matrice GLCM

Il existe différentes méthodes d'analyse de la texture des images. La première consiste en l'analyse des relations texturales à partir d'une fenêtre mobile se déplaçant de pixel en pixel sur l'image. Cette méthode présente certaines limitations, inhérentes à la méthode même. En effet, l'efficacité de l'utilisation d'une fenêtre mobile pour observer la texture d'une image varie selon la taille de la fenêtre. Une fenêtre de grande dimension permettra d'identifier de manière fiable les relations spatiales entre chacun des pixels avec pour principal défaut une diminution de la résolution de l'image résultante. A contrario, une fenêtre de faible dimension, générera des résultats à partir d'un faible nombre de paires de pixels, ce qui diminue la capacité de la fenêtre à capturer l'essence même des informations texturales. Cette méthode est également plus susceptible d'être sujette aux anomalies texturales.

La seconde approche est basée sur une analyse texturale par objet, après segmentation de l'image. Selon cette méthode, l'analyse texturale est réalisée à partir d'un plus grand nombre de paires de pixels. Les statistiques GLCM extraites de la méthode par objet sont considérées comme plus performantes pour la réalisation de classifications automatisées que les statistiques résultant de la méthode des fenêtres mobiles (van der Sanden and Hoekman, 1999).

Plusieurs études (Oliver, 2000; van der Sanden et Hoekman, 1999; Grover et al., 1999) démontrent que l'apport des variables texturales permet d'améliorer l'exactitude des classifications sur les écosystèmes forestiers modifiés par les exploitations. van der Sanden et Hoekman (1999) indiquent que, des neuf caractéristiques texturales présentées au Tableau 2, les variables texturales du contraste et de la corrélation, se sont montrées plus sensibles à la discrimination des écosystèmes forestiers. Ces dernières, contrairement aux autres statistiques GLCM, reflètent les propriétés biophysiques observées de la surface de la canopée (van der Sanden et Hoekman, 1999). La variable texturale du contraste s'est révélée efficace à la discrimination des forêts primaires vierges (présentant un contraste fort), les éclaircies (faible contraste) et les parcelles en voie de régénération (contraste intermédiaire) (Oliver, 2000).

1.7 Classification

La classification est définie comme le processus de catégorisation de pixels ou d'objets sous un ensemble de types d'occupations du sol ou plus généralement, de labels. Ci-dessous, une description de trois méthodes de classification couramment utilisées dans l'imagerie radar (Malik et Malhotra, 2016).

1.7.1 *Random Forest (RF)*

Ce classifieur est un algorithme non paramétrique d'apprentissage automatique développé par Breiman (2001) bien que Tin Kam Ho (1995) fut la première personne à introduire le concept de forêt de décision aléatoire. Cet algorithme construit plusieurs arbres de décisions sur un même jeu de données basé sur la technique du *bootstrap*, une méthode d'inférence statistique permettant d'approcher par simulation la distribution d'un estimateur lorsqu'on ne connaît pas la loi de l'échantillon (Efron, 1982). Cet algorithme a pour avantage de ne pas être influencé par le problème du surapprentissage survenant lorsqu'un modèle se base trop sur les données d'entraînement. Le phénomène du surapprentissage a pour conséquence que le modèle utilise le bruit et autres variabilités aléatoires de l'image pour son entraînement (Numbisi et al., 2018).

La première étape du lancement de l'algorithme est la génération d'arbres de décision pour plusieurs sous-échantillons réalisés avec remise. Les nœuds de chaque arbre de décision sont ensuite séparés selon une variable sélectionnant l'arbre présentant le plus haut taux de diminution d'impuretés (Hastie et al., 2009).

Cet algorithme est fréquemment utilisé pour les classifications d'images SAR, que ce soit dans les zones forestières (Balzter et al. 2015; Rüetschi et al., 2018) ou culturelles (Ghazaryan et al., 2018). Bien que dans certains cas, l'algorithme *Random Forest* soit toujours soumis à des problèmes de surapprentissage ayant pour résultat des classifications de faible qualité (Zhao et al., 2018), il est considéré comme plus efficient par rapport à d'autres méthodes, notamment pour la classification de cultures (Ghazaryan et al., 2018) et de mangroves (Abdel-Hamid et al., 2018 ; Numbisi et al., 2018).

1.7.2 *Support vector machine (SVM)*

Le principe du SVM est de définir un hyperplan permettant la séparation de pixels ou d'objets appartenant à des classes différentes. Cette séparation peut s'effectuer de manière linéaire ou non. Dans le cas des séparations non linéaires, deux possibilités peuvent être envisagées. La première consiste en la création d'une marge souple pouvant s'adapter aux données présentant beaucoup de bruits. La seconde s'appuie sur l'utilisation de la fonction de Kernel. Cette fonction simule la projection du jeu de données d'intérêt dans un espace présentant un nombre supérieur de dimensions que celui du jeu de données initial. Dans ce nouvel espace créé, les données sont séparées linéairement (Lardeux et al., 2006 ; Malik et Malhotra, 2016).

Cette définition des SVM s'applique bien aux cas de séparations binaires. Cependant, la majorité des jeux de données utilisés pour la télédétection satellitaire dépassent la simple dimension binaire. Afin de répondre à ce problème de multiplication des classes, deux variantes ont été développées (Lardeux et al., 2006).

1.7.2.1 One Against All (1AA)

Cet algorithme, développé par Vapnik (1995), consiste en la construction de n hyperplans permettant une séparation effective d'une classe particulière suivie par les $n-1$ autres classes (Lardeux et al., 2006). Un point serait donc accepté par une classe si et seulement si le SVM de la classe en question accepte ce point et qu'il est rejeté par le SVM de chacune des autres classes (Aisen, 2006).

1.7.2.2 One Against One – Classification par pair (1A1)

Cette méthode utilise un SVM par paire de classes nécessitant une séparation résultant en la création de $N(N-1)/2$ SVM. Une fois appliquée à un point, chaque classification confère un vote à la classe gagnante et le point est finalement labellisé selon la classe ayant collecté le plus de votes. Cette approche simpliste peut être modifiée et améliorée par pondération de chacune des classes (Gidudu et al., 2014).

Afin d'en identifier leur efficacité respective, Gidudu et al. (2014) évaluent les méthodes 1AA et 1A1 par comparaison de quatre classifications de couverture du sol, provenant de données Landsat. Ces quatre classifications sont issues de l'application de la fonction de Kernel sous une forme linéaire, quadratique, polynomiale et selon la fonction radiale de base (RBF). Il fut conclu que la forme linéaire de la fonction de Kernel ainsi que la fonction RBF présentaient de plus faibles résultats de classification que les formes polynomiales et quadratiques. Cette différence est cependant considérée comme négligeable d'un point de vue statistique selon un intervalle de confiance de 95%. Les auteurs concluent leur article en indiquant que le choix d'une méthode ou d'une autre se fait selon les préférences de l'utilisateur.

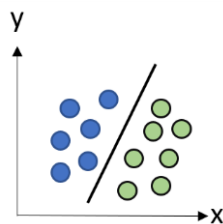


Figure 8 : Classification linéaire en utilisant un SVM (Malik and Malhotra, 2016)

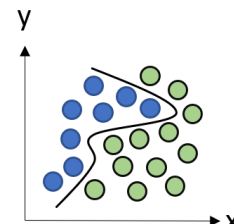


Figure 9 : Classification non linéaire en utilisant un SVM (Malik and Malhotra, 2016)

1.7.3 Analyse discriminante linéaire (LDA) – Réduction de dimension

L'analyse discriminante linéaire est une généralisation de la méthode de classification de Fischer. La méthode de Fischer cherche à identifier les combinaisons linéaires des éléments caractérisant ou séparant deux ou plusieurs classes d'objets. La combinaison résultante peut être utilisée en tant que classifieur linéaire ou pour réduire le nombre de dimensions d'un jeu de données en vue d'une future classification. La projection des données dans ce nouvel espace à une dimension est réalisée de manière à maximiser la différence entre les moyennes des différentes classes d'objets tout en minimisant la variance intra-classe (Malik et Malhotra, 2016; Brown, 1999). L'utilisation de cette méthode permet par le passé la classification de milieux variés allant des zones agricoles (Sakarya et Demirpolat, 2018) aux icebergs (Bogdanov et al., 2005).

1.8 Conclusion de la revue de la littérature

Comprises dans des latitudes inférieures à 30° nord et sud de l'équateur, les forêts tropicales couvrent 13,5 % des terres émergées. Ces dernières, bien qu'elles fassent déjà l'objet de nombreuses études, sont particulièrement difficiles à caractériser. Cette difficulté s'explique principalement par leur dynamique naturelle, leur constitution complexe ainsi que leur accessibilité limitée. La télédétection satellitaire est une méthode d'acquisition des données particulièrement pertinente pour étudier les forêts tropicales car celle-ci permet de répondre à ce problème d'accessibilité.

L'usage des données radar dans la cartographie satellitaire n'est pas chose récente. Elles présentent de nombreux avantages par rapport aux données optiques tels que leur capacité à ignorer les effets atmosphériques ou à pouvoir acquérir des données durant le jour comme la nuit. L'instrument SAR se montre dès lors particulièrement intéressant dans les zones soumises à d'importantes dégradations atmosphériques tels que les milieux tropicaux.

Les écosystèmes forestiers tropicaux sont des entités complexes soumises à de nombreuses transformations, qu'elles soient liées aux multiples acteurs associés à ces milieux ou aux changements climatiques globaux. L'évolution de la technologie radar est donc prometteuse dans le sens où elle permettrait de caractériser et de réaliser des suivis de ces milieux.

L'analyse polarimétrique des données radar combine régulièrement les données des coefficients de rétrodiffusion avec les données texturales, ces dernières permettant d'améliorer les classifications. Deux méthodes de classification sont utilisées pour la discrimination des écosystèmes forestiers. La première, basée sur l'extraction des valeurs des pixels, présente des performances modérées suite à la contamination des images par le speckle. Il existe toutefois de nombreux filtres permettant de limiter l'effet de ce dernier.

La seconde méthode, basée sur l'extraction des valeurs de réflectance et de texture par objets, nécessite une segmentation de l'image. Dans ce cas, les algorithmes MRS, SLIC et SLICO se montrent particulièrement performants pour réaliser des segmentations des images radar des satellites S1. Les opérations de classification des images peuvent également être réalisées sur bases de plusieurs algorithmes. L'algorithme *Random Forest* est couramment utilisé dans la discrimination des écosystèmes forestiers pour ses performances générales de classification, mais également pour sa capacité à éviter le surentrainement. Cette dernière permet de réaliser des classifications par pixel de qualité, et ce, en limitant l'impact du speckle.

En considération de ces différentes possibilités, plusieurs traitements sont à tester afin d'identifier la capacité du radar à réaliser une discrimination des écosystèmes forestiers du bassin du Congo.

1.9 Objectifs

A la demande de l'Observatoire des forêts d'Afrique centrale et de la Commission des forêts d'Afrique centrale (COMIFAC), l'objectif de cette étude est de développer une méthode de traitement des séries temporelles des satellites S1 permettant la cartographie des écosystèmes forestiers de la République du Congo et selon l'efficacité de la méthode, à une plus large échelle pouvant s'étendre sur tout le bassin du Congo.

De nombreuses études ont déjà été réalisées par le passé afin de cartographier les forêts tropicales. Les technologies liées au domaine de la télédétection ne cessent de s'améliorer et permettent de réaliser des classifications de plus en plus fines des écosystèmes forestiers. Ces avancées technologiques ont permis de passer des classifications binaires forêt/non forêt, à la discrimination des forêts sèches et marécageuses jusqu'à l'identification des écosystèmes forestiers.

Par le passé, ces classifications ont majoritairement été réalisées sur base de données optiques ou en combinant des données optiques et radar ; les données radar permettant d'améliorer les résultats générés à partir des données optiques. S'affranchir des données optiques permettrait la cartographie de nombreux milieux où l'acquisition de données optiques est impossible à cause des phénomènes atmosphériques.

Outre leur résolution spatiale importante (10 m) et contrairement aux données SAR antérieures, les données des satellites S1 ont pour avantage d'être abondantes (une image tous les 6 jours en Europe et tous les 12 jours pour le reste du monde) et gratuites. Dans le cas où la seule utilisation des données S1 permettrait d'effectuer une discrimination précise des écosystèmes forestiers tropicaux, l'abondance d'image des satellites S1 offre l'opportunité de réaliser des mises à jour régulières ainsi qu'un suivi pour évaluer l'évolution de ces milieux.

Ce mémoire s'inscrit donc dans la continuité de ce souhait d'affinement de discrimination des écosystèmes forestiers tropicaux tout en s'affranchissant de l'utilisation des données optiques pour se concentrer uniquement sur l'utilisation des données des satellites Sentinel-1 (SAR-C) et ALOS-2 (PALSAR-2). Il permettra dès lors d'évaluer la capacité des images radars à contribuer à une telle discrimination.

Chapitre 2 : Zone d'étude

Le travail sur le terrain effectué dans le cadre de ce mémoire s'est réalisé dans le Nord de la République du Congo, sur les concessions forestières de la Congolaise Industrielle des Bois (CIB). La zone d'étude couvre une surface totale de 6.600 km² et s'étend de la position nord-ouest (16°3'5 446'' E, 2°13'34 385''N) au sud-est (16°45'19 706''E, 1°6'27 741''). La CIB est une société industrielle et commerciale spécialisée dans le secteur d'activité de la gestion forestière, de l'exploitation, de la transformation et de la commercialisation de bois tropicaux. Cette société, pionnière en matière de gestion durable des forêts tropicales, appartient au groupe Olam, un des leaders mondiaux dans la gestion de la chaîne d'approvisionnement de matières premières et produits agricoles (PFBC, 2012). Afin d'appuyer ce souhait d'exploitation durable des forêts, la CIB a certifié trois de ses quatre concessions forestières sous le label FSC. Un plan d'aménagement visant à certifier la quatrième concession est en cours de validation par les autorités locales. Une telle certification vise à améliorer la production du bois de manière à ce que celle-ci soit écologiquement appropriée, socialement bénéfique et économiquement viable (FSC, s.d). La certification de ces différentes concessions, en combinaison avec une production de bois élevée à 395.000 m³ en 2017, confère à la CIB la première place des entreprises d'exploitation forestière en milieu tropical certifiée FSC au monde.

Les concessions forestières de la CIB, de par la présence d'un réseau routier servant à la circulation des hommes et au transport des grumes, permettent de répondre à la principale difficulté que l'on rencontre en forêt tropicale : l'accessibilité. La facilitation de la mobilité dans ces milieux représente une opportunité unique pour réaliser des campagnes de collecte de données.

2.1 Climat

Le botaniste Wladimir Peter Köppen inventa en 1900 une classification empirique des climats basée sur des paramètres de température et de pluviométrie (Arnfield, 2019). Un tel découpage a permis la création de 5 grands types climatiques ainsi que de 30 sous-types.

Afin d'en améliorer sa justesse, cette classification fut revisitée plusieurs fois par le passé (Beck et al., 2018; Peel et al., 2007). Sur base de ces récentes classifications, la zone d'étude est située à cheval sur deux types de climats. Le premier correspond à un climat tropical de forêt humide, symbolisé par l'association des lettres *Af*. Le second correspond, quant à lui, à un climat tropical de mousson, symbolisé par l'association des lettres *Am*.

La lettre *A*, représentative des climats tropicaux, indique que le mois le plus froid de l'année doit être supérieur à 18 °C. La lettre *f* indique que les précipitations du mois le plus sec doivent être supérieures à 60 mm. Finalement, la lettre *m*, indique que les précipitations du mois le plus sec sont inférieures à 60 mm, mais peuvent être égale ou supérieure à 100 mm – $MAP/25$; MAP correspondant à la moyenne des précipitations annuelles.

Chapitre 2 : Zone d'étude

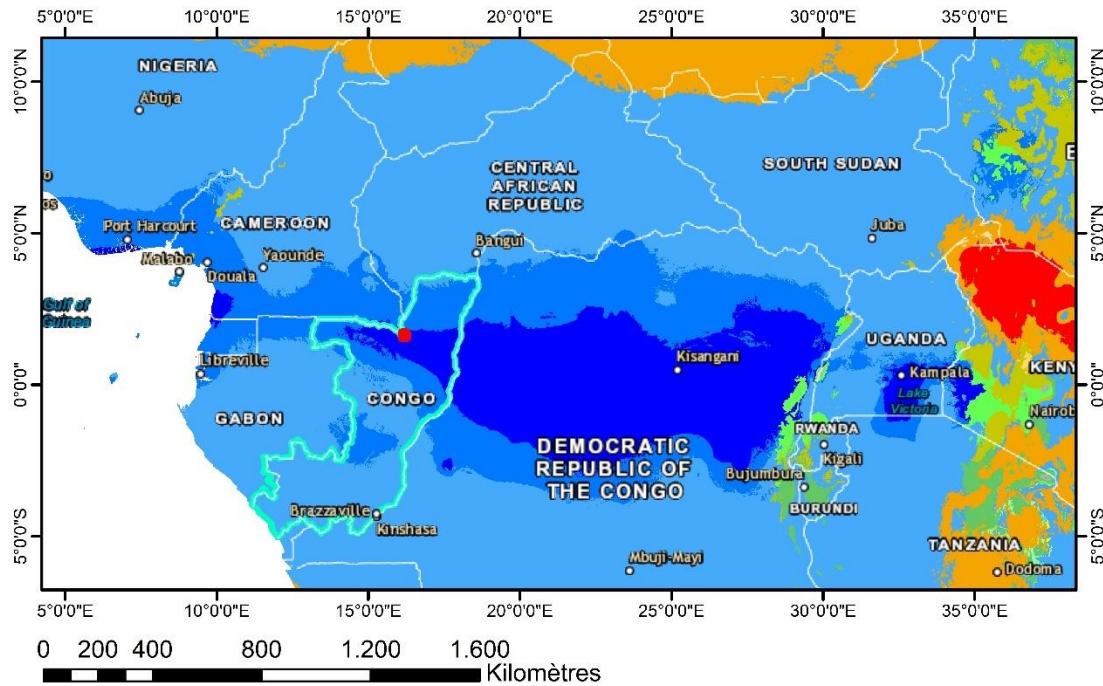


Figure 10 : Classification des climats de Köppen-Geiger révisée par Beck et al., 2018.



Le climat de cette zone présente deux saisons humides et deux saisons sèches (Samba et al., 2008). La première saison sèche, la plus longue, s'étend de décembre à fin février. La seconde, plus courte, s'étend de fin juin à fin juillet. Les précipitations moyennes annuelles sont de 1.800 mm en 115 jours (Bulltot, 1974).

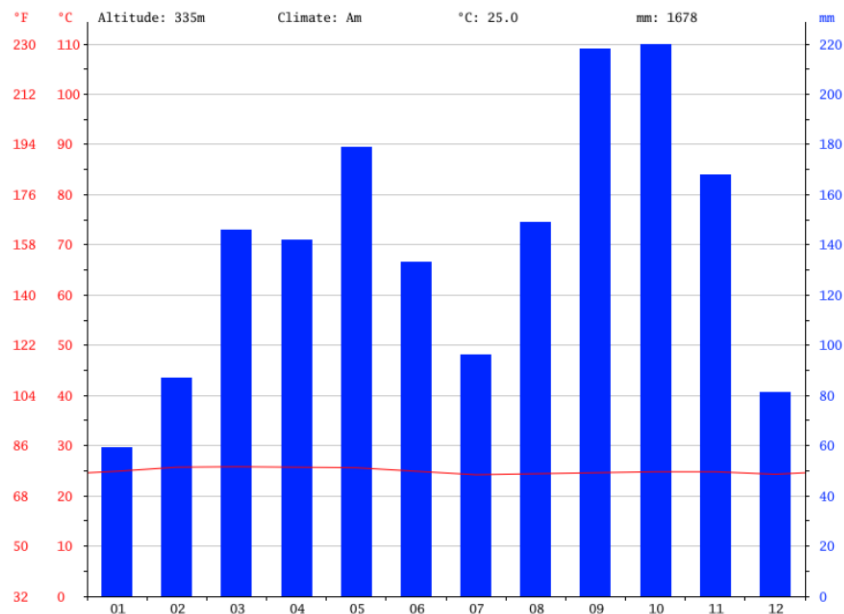
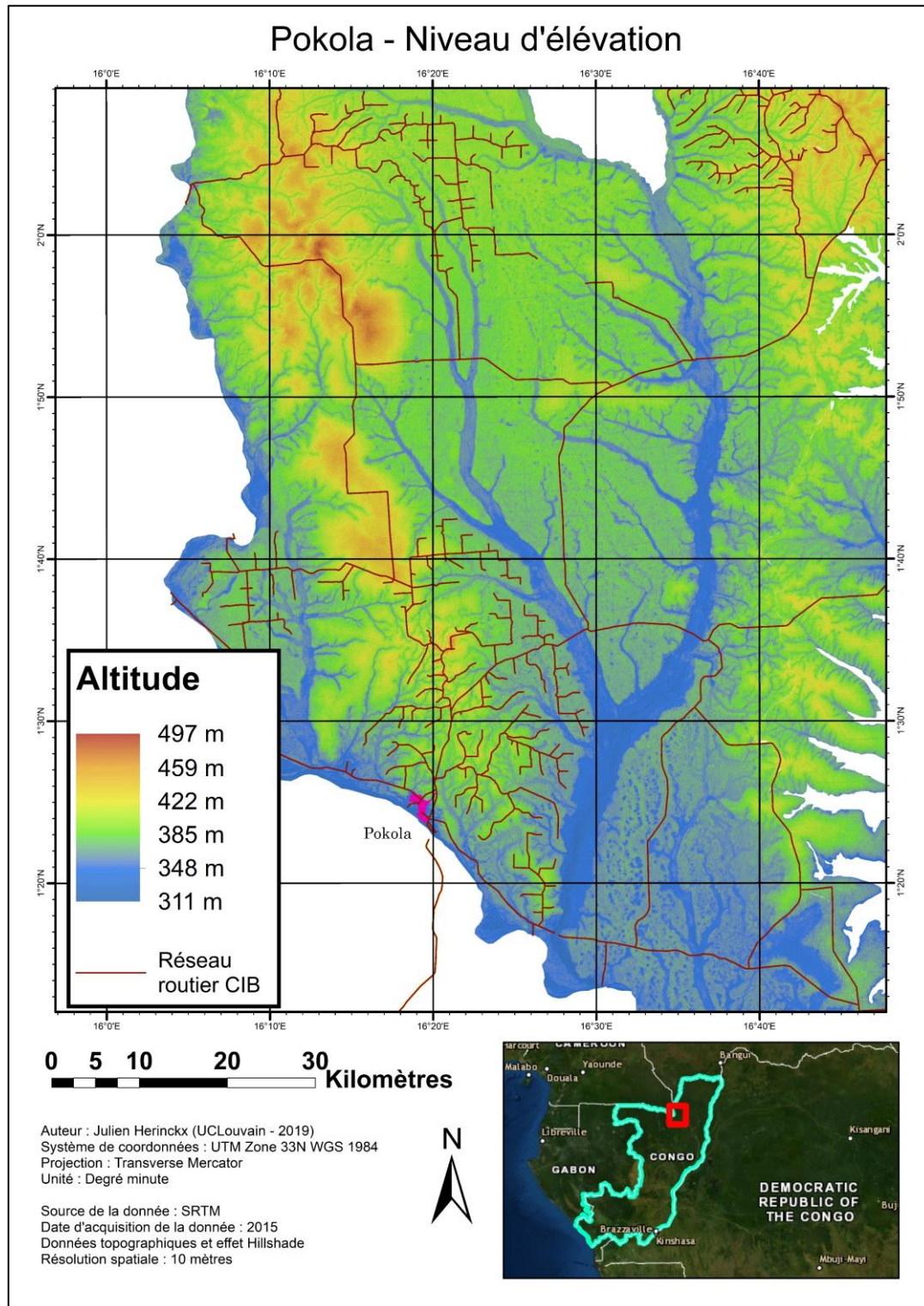


Figure 11: Diagramme climatique de Pokola (<https://fr.climate-data.org/afrique/congo-brazzaville/sangha/pokola-52835/>)

2.2 Topographie

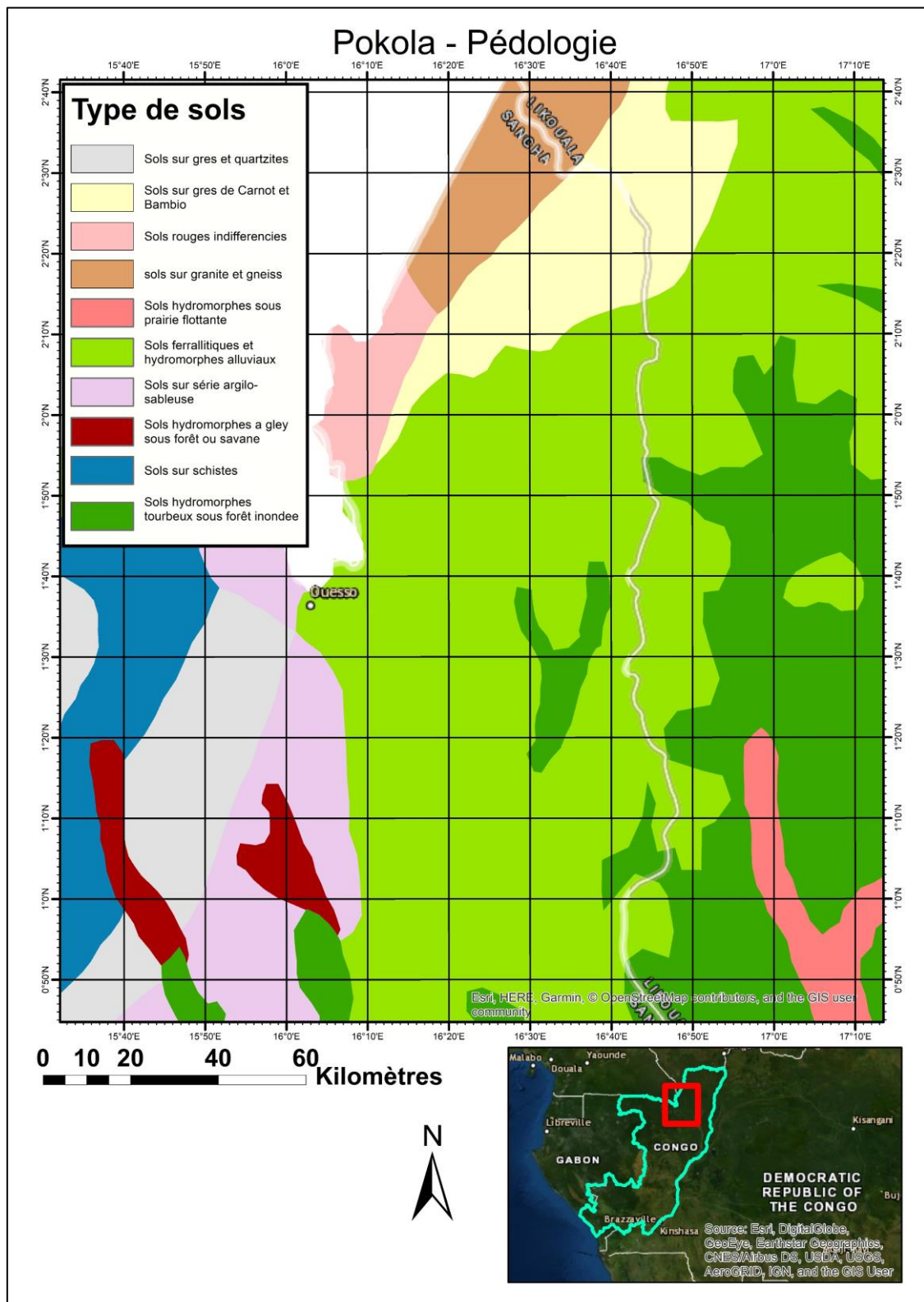
La zone d'étude présente une altitude moyenne de 385 mètres. Elle est bordée d'une part, par le fleuve de la Sangha qui coule sur tout le flanc ouest des concessions et d'autre part, par les nombreuses formations marécageuses limitant les déplacements sur toute la partie est. Bien que sillonnée par de nombreuses zones marécageuses et cours d'eau, la topographie de la région est relativement homogène.



Carte 2 : Topographie des concessions forestières de la CIB.

2.3 Pédologie

Pokola et ses environs sont situés sur des sols ferrallitiques. Les zones marécageuses se caractérisent quant à elles par des sols hydromorphes tourbeux.



Carte 3 : Données pédologiques des sols des concessions forestières de la CIB. Données pédologiques mises à disposition par la cellule d'aménagement de la CIB.

Chapitre 3 : Matériel et méthode

Ce chapitre sera découpé en trois sections :

La première d'entre elle vise à décrire l'ensemble des données utilisées dans le cadre de cette étude ainsi que les prétraitements qui ont été appliqués à ces données. Cette section sera clôturée par la description des données acquises lors de la campagne de terrain. Etant étroitement liées à la manière avec laquelle elles ont été collectées, ces données seront décrites en parallèle du protocole de terrain.

La seconde section décrira, sur base du matériel présenté à la précédente section, les analyses préliminaires permettant de cibler la méthodologie la plus pertinente à appliquer pour le traitement des données. Trois éléments seront ainsi analysés : la phénologie du milieu, les textures de l'image permettant la discrimination des forêts primaires et des forêts secondaires ainsi que les performances de différents algorithmes de segmentation.

La troisième et dernière section présentera l'ensemble de la méthodologie appliquée aux images. Celle-ci couvrant les opérations de traitements, les opérations de segmentation et de classification, la validation de ces dernières ainsi que les opérations de post-classification.

3.1 Matériel

3.1.1 Acquisition des données – SAR-C

Les données SAR-C ont été acquises par le satellite Sentinel-1A depuis l'orbite descendante relative 7 (orbite absolue 20029), et ce, sur la période du 01/01/2018 au 26/12/2018. Elles furent acquises sur une fauchée de 250 km de large, selon le mode IW avec deux polarisations : VV et VH.

Les données distribuées font l'objet d'un prétraitement variant selon le produit désiré. Les données utilisées pour cette étude ont connu le traitement *Single Look complexe (Level-1 SLC)*. Ce type de donnée contient les informations originales sur l'amplitude et la phase capturées par l'instrument SAR. Ces données font également l'objet d'une correction sur les erreurs de réception que le satellite a pu connaître lors de l'acquisition, notamment sur les erreurs de position sur la latitude et la longitude (ESA, 2019). Afin d'extraire un maximum d'information des polarisations VV et VH, une composition colorée fut créée tenant compte des informations fournies par ces deux polarisations ainsi que le ratio de VV sur VH. Bien que ce ratio n'apporte aucune information brute, il permet de faciliter la lecture de l'image via la mise en évidence d'objets d'intérêt tels que les peuplements de limballis ou certaines forêts marécageuses.

Le mode IW, couplé avec la technique TOPSAR consiste en l'acquisition de données sur trois sous-fauchées. Chacune de ces sous-fauchées est définie par une

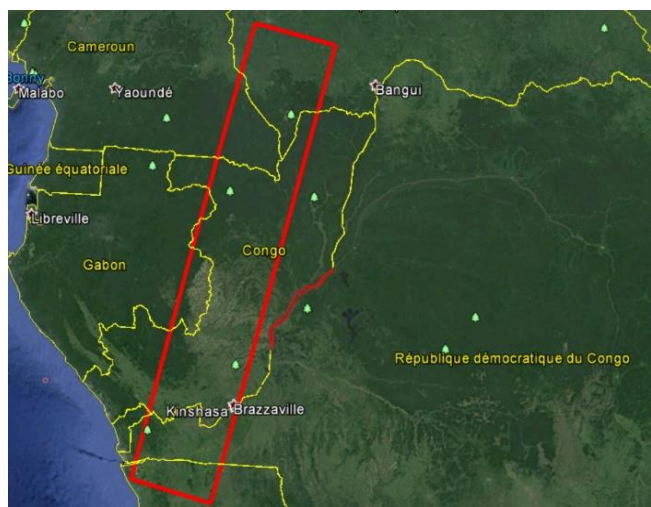


Figure 12 : La marque rouge de l'image correspond à l'emprise des données capturées par le satellite Sentinel-1A sur l'orbite absolue 20029. Echelle : 1 : 6 000 000. Image de fond SIO, NOAA, U.S. Navy, NGA, GEBCO (Google Earth).

gamme d'angles d'incidence limite ne dépassant pas 5 degrés par sous-fauchée de manière à limiter les distorsions radiométriques. Les données concernant cette étude ont été acquises sur la troisième sous-fauchée du passage du satellite. Le Tableau 3 renseigne sur les différents angles d'incidence de ces sous-fauchées.

Tableau 3 : Angle d'incidence des données acquises selon le mode IW. La troisième sous fauchée (IW3) correspond aux données utilisées pour la zone d'étude.

Sous-fauchée	Angle d'incidence minimal	Angle d'incidence maximal
IW1	30,07°	36,64°
IW2	36,40°	41,88°
IW3	41,68°	46,06°

3.1.1.1Prétraitement

Les données brutes n'étant pas directement exploitables en vue d'une analyse polarimétrique, plusieurs prétraitements doivent être réalisés.

Orbitale

Les informations sur l'orbitale contenues dans les métadonnées des produits Sentinel-1 manquent généralement de précision. Cette précision peut être affinée grâce aux informations sur la position exacte du satellite ainsi que sa vitesse. Ces éphémérides, plus précises, sont recalculées a posteriori et diffusées quelques semaines après l'acquisition de l'image. Elles permettent de mettre à jour les informations contenues dans les métadonnées de l'image d'origine.

Calibration

Cette étape de calibration radiométrique est nécessaire pour convertir les données SAR brutes en mesures quantitatives. Cette conversion permet de réaliser des comparaisons entre des données SAR acquises par différents capteurs. Les données mises à disposition depuis le *Copernicus Open Access Hub* ont connu le traitement Level-1 SLC, ce dernier n'incluant pas de calibrations radiométriques. En conséquence, de nombreux biais non négligeables sont présents sur ces données collectées. Cette étape permet donc de réaliser l'exacte correspondance entre la valeur d'un pixel donné et le coefficient de rétrodiffusion de la surface réfléchissante qui lui est liée.

Correction du bruit d'origine thermique

L'acquisition de données radar, en particulier pour les polarisations croisées HV et VH, présente des erreurs radiométriques liées à des modifications de température du capteur ainsi qu'à l'émissivité de la terre même. Cette correction permet de limiter le bruit inhérent aux modifications de températures du capteur (Piantanida, 2017).

Aplatissement de terrain

Lorsqu'une classification de la couverture du sol est appliquée à un terrain qui présente des variations topographiques, des erreurs peuvent être générées. Ces effets topographiques peuvent affecter la position d'un objet d'intérêt, mais également modifier la luminance de la surface étudiée captée par le radar. Cet opérateur permet donc de supprimer la variabilité radiométrique associée à la topographie en utilisant la méthode d'aplatissement de terrain proposée par Small, et ce tout en conservant la variabilité radiométrique associée à la couverture terrestre (Small, 2011).

Correction de terrain

Les variations topographiques ainsi que l'angle d'incidence local peuvent causer des distorsions des distances. Les données acquises sur les surfaces dont la position n'est pas située directement au Nadir du capteur seront toutes sujettes à de telles distorsions. Les données acquises par le capteur SAR-C du satellite S1 présentent, dans le cas du mode d'acquisition IW, un angle d'incidence minimal de 30,07°, nécessitant une correction de terrain. Cet opérateur permet de limiter et de corriger les distorsions de manière à ce que la géométrie représentée à l'image corresponde au mieux à celle de la surface réelle.

3.1.2 Acquisition des données – PALSAR

Les données utilisées dans le cadre de cette étude sont des mosaïques, dont les coefficients de rétrodiffusion, acquis par l'instrument PALSAR-2 du satellite ALOS-2, ont été moyennés sur l'année 2017 (dernières données en date disponibles). Ces mosaïques contiennent des données sur les polarisations HH et HV à une résolution de 25 mètres. Ces données ont fait l'objet de corrections permettant de limiter les effets topographiques (correction des pentes) ainsi que les distorsions géométriques (correction par orthorectification) (Shimada et al., 2014). Les données sont finalement normalisées sur une échelle 8-bit. Le seul traitement supplémentaire appliqué à ces données fut le passage d'une fenêtre mobile de taille 5X5 pixels permettant l'application du filtre antispeckle de Lee.

3.1.3 Protocole et observations de terrain

Un échantillonnage de données fut réalisé en République du Congo, à proximité de Pokola, sur les concessions forestières de la Congolaise Industrielle des Bois (CIB). Cette campagne de terrain permet l'identification de plusieurs écosystèmes et associations végétales types par l'intermédiaire d'une collecte de données sur 52 zones distinctes. Le travail de terrain fut découpé en deux étapes :

La première de ces étapes consistait en l'analyse préliminaire d'images radar et optiques provenant des satellites S1, ALOS et S2. Ces images n'ont fait l'objet que de prétraitement sommaire ayant toutefois permis l'identification de multiples zones d'intérêt. Le Tableau 4 renseigne sur les différentes compositions colorées utilisées pour réaliser l'identification des milieux d'intérêt. Une fois les différents milieux identifiés, une sélection de ces derniers fut réalisée sur base de leur homogénéité d'un point de vue spectrale ainsi que de leur facilité d'accès. La mobilité pédestre étant très limitée en forêt, il est nécessaire d'optimiser chacun des déplacements. Cette affirmation est également valable pour les déplacements en voiture. En considération de ces différents paramètres, entre 3 et 5 zones ont pu être échantillonnées par jour de collecte de données sur le terrain.

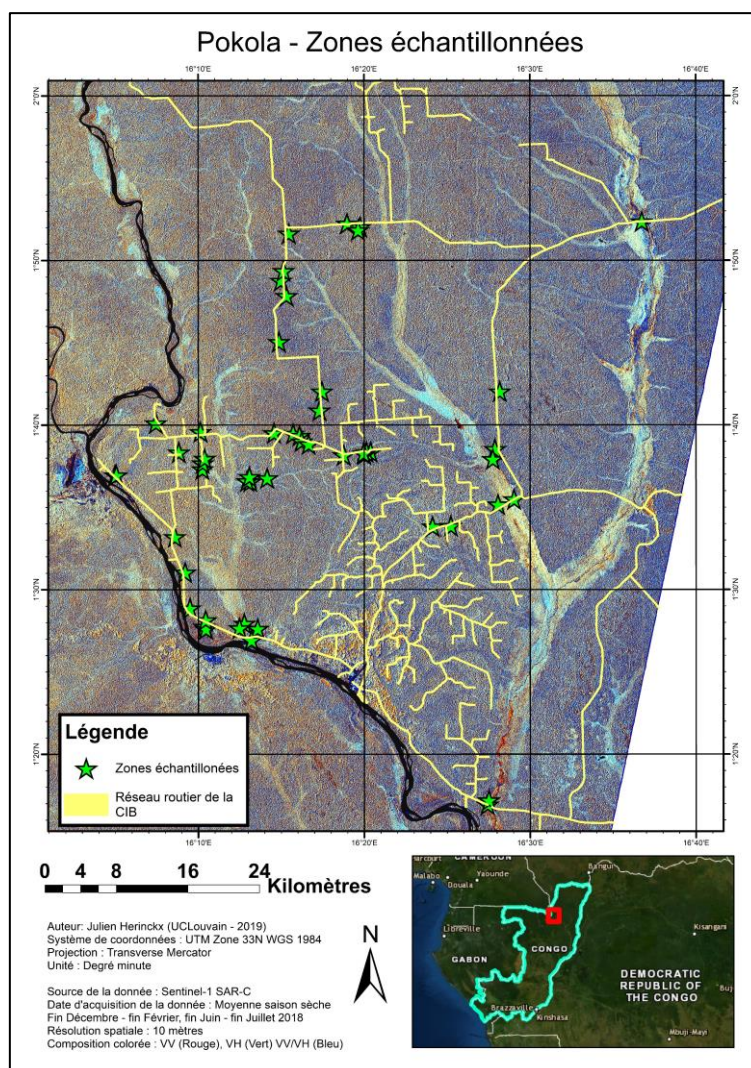
Tableau 4 : Ensemble des images satellites ayant permis l'identification des zones d'intérêt lors de la campagne de terrain. Les données S2 ont uniquement été utilisées dans le cadre de cette campagne afin de faciliter la différenciation des écosystèmes forestiers, la composition [RedEdge, SWIR2, NIR] permettant d'apporter un maximum de contraste entre les différentes formations forestières.

Satellite	Instrument	Composition colorée (RGB)
Sentinel-1	SAR-C	VV, VH, VV/VH
ALOS-2	PALSAR-2	HH, HV, HH/HV
Sentinel-2	Imageur multispectral (MSI)	RedEdge, SWIR2, NIR

La seconde étape consista en la récolte de données sur terrain. Pour ce, faire des placettes carrées de 2500 m², tracées à partir d'un GPS Garmin MAP62st, ont été installées dans chacune des zones à analyser afin d'y étudier cinq paramètres majeurs : l'étagement de la

végétation, les essences en présence, la présence de chablis et trouées ainsi que la rugosité de la canopée. Une telle dimension de placette permet d’acquérir des données sur une surface suffisamment grande que pour confirmer l’homogénéité du milieu. Cela permet également d’associer les informations de terrain à la signature spectrale du milieu correspondant de manière à pouvoir dégager les variables responsables d’un changement de signal. Les données liées à l’identification des essences furent réalisées grâce à un botaniste expert en la reconnaissance d’essences tropicales.

Les concessions forestières sont traversées par un réseau de pistes, nécessaire au transport des hommes, grumes et outils. L’installation d’un tel réseau a bien évidemment eu des impacts considérables sur les dynamiques naturelles des successions végétales des écosystèmes situés à proximité des pistes. Dans l’optique de s’affranchir de l’influence de ce réseau, les placettes furent installées à une distance minimale de 200 mètres des bords des routes. Les déplacements vers chacune des zones d’intérêts ont été réalisés en accompagnant les ouvriers de la CIB vers les lieux d’exploitation ou en utilisant les véhicules gracieusement mis à disposition par la CIB. Les déplacements doux furent facilités par un machetteur, ouvrant l’accès à la zone d’intérêt. L’ensemble des données acquises durant le travail sur le terrain est reproduit en annexe.



Carte 4 : Zones échantillonnées lors de la campagne de terrain. Au total 52 échantillons furent collectés, à raison d’un échantillon par zone (symbolisée par l’étoile verte).

Etagement de la végétation

Elle peut être découpée en quatre strates principales :

- **Strate émergente** : composée d'arbres dominants dont la hauteur est significativement supérieure à la moyenne des arbres en présence. Ces arbres peuvent être présents de manière ponctuelle ou former un tissu discontinu.
- **Strate co-dominante** : composée d'arbres dont la hauteur individuelle est égale à la hauteur moyenne des autres arbres de cette strate. Cette strate forme généralement un tissu continu.
- **Strate dominée** : composée d'arbres dont la hauteur est strictement inférieure à celle des arbres dominants ou co-dominants.
- **Strate arbustive** : composée de plantes ligneuses de dimension inférieure à 12 mètres (Jennings et al., 1999).

Les zones ouvertes ou inondées ne présentant pas d'essences ligneuses ont été caractérisées par une strate herbacée. La Figure 13 reprend une description visuelle de ces différentes strates. Afin de quantifier l'importance d'une strate par rapport à une autre, le pourcentage de recouvrement de chacune d'entre elles fut quantifié par estimation visuelle, en concertation avec l'expert botaniste. Le pourcentage de recouvrement correspond à la surface projetée au sol de l'ensemble des houppiers de la strate d'intérêt. Cette surface est ensuite rapportée à la surface totale de la placette (2500 m²). La somme de ces pourcentages peut donc être supérieure à 1, dans ce cas, les houppiers des arbres de strates différentes se recouvrent.

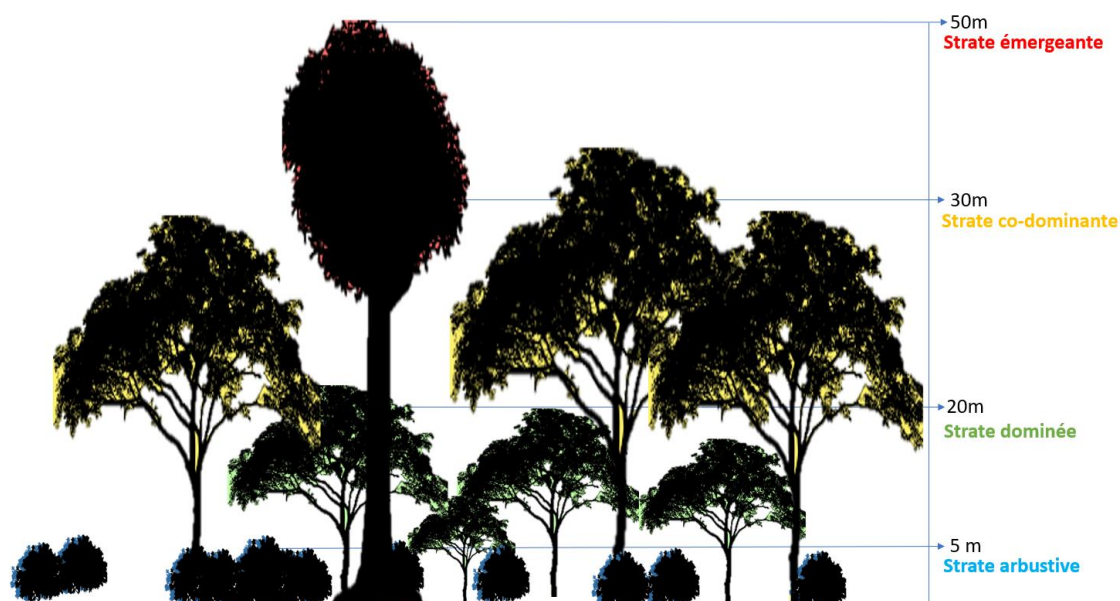


Figure 13 : Interprétation visuelle des quatre principales strates rencontrées sur le terrain. Les mesures de hauteur [mètres] ne sont présentes qu'à titre indicatif et font référence aux cas les plus fréquemment rencontrés.

Phénologie

Afin de pouvoir étoffer les données sur la phénologie du milieu, de multiples informations liées aux variations saisonnières des propriétés hydrologiques de chacune des zones ainsi qu'à la sénescence des feuilles des différentes essences furent collectées. Cependant, la collecte de données sur la sénescence du feuillage ne permit de tirer aucune conclusion, les savoirs locaux ne permettant pas toujours de répondre à une telle question.

Essence en présence

Les essences majoritaires composant chacun des milieux ont été identifiées à l'aide d'un botaniste expert en la reconnaissance d'essences tropicales. Leur densité respective ainsi que leur strate fut répertoriée pour chacune des zones visitées, dans la mesure du possible. Les données liées à la densité de chacune des essences ont été acquises par estimation du pourcentage de recouvrement, en se basant sur le coefficient de Braun-Blanquet.

Tableau 5 : Comparaison du coefficient de Braun-Blanquet avec le pourcentage de recouvrement d'une essence d'intérêt par projection de la surface de son houppier au sol.

Coefficient de Braun-Blanquet	Pourcentage de recouvrement
+	Moins de 10 individus
1	Moins de 5 %
2	5 – 25 %
3	25 – 50%
4	50 – 75%
5	Plus de 75 %

Chablis et trouées

Le nombre de chablis ainsi que leurs dimensions furent répertoriés sur les placettes de chacune des zones d'intérêt. Selon une perspective spectrale, la présence de chablis et trouées influe sur les propriétés texturales de la zone. Par exemple, une zone non exploitée présentera une structure à maturité homogène (Figure 14) alors qu'une zone exploitée connaîtra l'alternance entre des arbres de grandes dimensions et de nombreuses trouées. La canopée, dans ce cas, sera très irrégulière d'un point de vue textural (Figure 15).



Figure 14 : Les zones non exploitées, lorsqu'elles ne sont soumises à aucune autre perturbation naturelle, présentent généralement une structure climacique, la canopée y est homogène et fermée. Certaines formations forestières peuvent, au contraire des structures mono-climaciques, présenter une structure méta-climacique. Les stades climaciques se succèdent induisant une canopée pouvant être hétérogène.



Figure 15 : Les zones exploitées présentent de fortes variations de hauteur sur de faibles distances induisant une texture particulière à l'image. Les essences en présence sont différentes de celles des structures climaciques et sont majoritairement pionnières à post-pionnières, selon la date de la perturbation

Topographie

Bien qu'une correction de terrain et de pente ait été appliquée sur les images SAR à l'aide d'un modèle de terrain, il est toujours pertinent de relever les informations topographiques locales, car elles peuvent être liées à des modifications du coefficient de rétrodiffusion. À l'exception des zones marécageuses, la topographie est très homogène et plane sur la quasi-totalité des zones de forêt parcourues.

Transect

Afin d'évaluer les propriétés texturales de ces zones, 42 transects d'une longueur comprise entre 200 et 500 mètres ont été tracés dans les zones d'intérêt. Dix transects des 52 zones initialement échantillonnées n'ont pas pu être tracés, les dix zones correspondantes ne présentant pas de couvert arboré ou étant inondées rendant difficile tout déplacement. Sur ces transects furent collectés des mesures de la hauteur maximale de la canopée, à raison d'une mesure tous les dix mètres. Les appareils à pointeur (Vertex, relascope de Bitterlich) sont particulièrement difficiles à manipuler en forêt tropicale, la canopée, trop densément fermée, empêche d'avoir une bonne visibilité du faite de l'arbre. En conséquence, les mesures de hauteur furent réalisées par estimation visuelle en concertation avec le botaniste. L'écart-type de ces différentes mesures de la hauteur sert d'indicateur pour évaluer la rugosité de la canopée.

Afin de mieux percevoir cette notion de rugosité, le lecteur est invité à prendre connaissance des Figure 16 et Figure 17. Ces figures représentent les données acquises sur deux transects, l'un en forêt secondaire, dont la dernière exploitation remonte à l'année 2014 et l'autre en forêt primaire. Les forêts sujettes aux récentes exploitations présentent un bien plus grand écart-type au niveau de la hauteur des arbres que les forêts primaires. Cette différence, bien que très visible sur le terrain, est difficilement identifiable sur les images radar en bande C (van der Sanden et Hoekman; 1999).

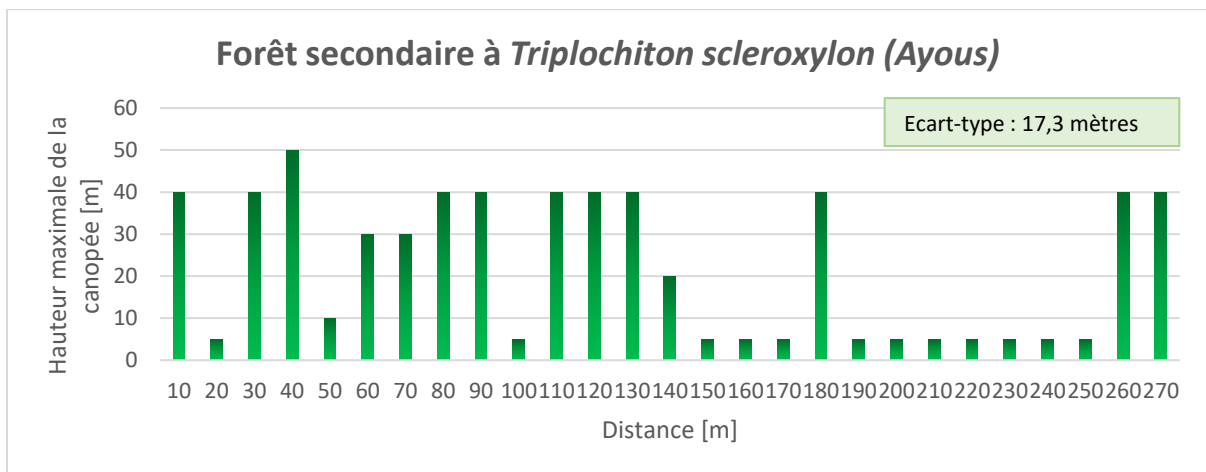


Figure 16 : Analyse de la hauteur maximale des arbres depuis un transect de 270 mètres en forêt secondaire à *Triplochiton scleroxylon* (Ayous) située à la position 16°24'7"E – 1°33'51"N

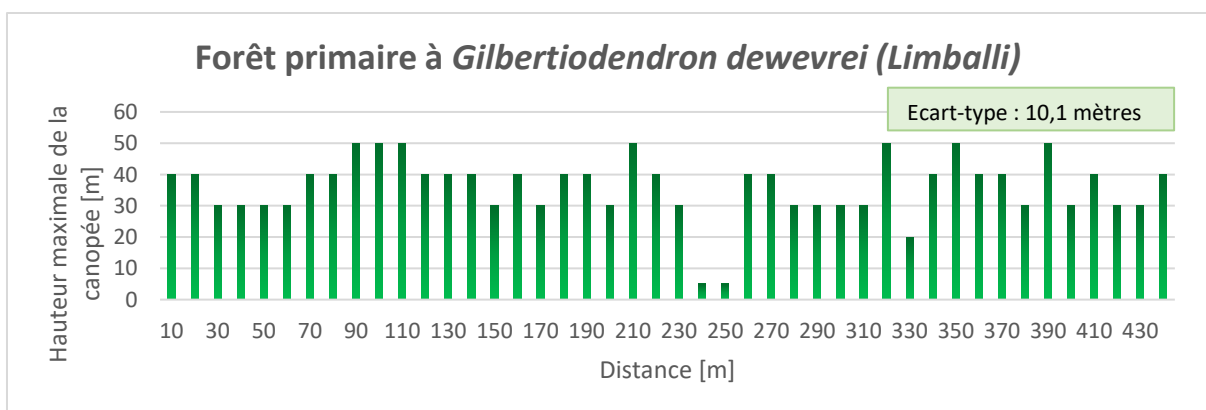


Figure 17 : Analyse de la hauteur maximale des arbres depuis un transect de 430 mètres en forêt primaire dominée par le *Gilbertiodendron dewevrei* (Limballi) située à la position 16°27'53"E – 1°38'33"N

3.2 Méthode

3.2.1 Analyse préliminaire

3.2.1.1 Phénologie

Les images collectées sur l'année 2018 furent moyennées en fonction des saisons de végétation. La Figure 18 reprend l'évolution de l'Indice de Végétation Normalisé (NDVI), au cours de l'année, des espaces forestiers se trouvant au nord de Pokola. Ces données sont acquises par le capteur MERIS. Elles correspondent à une moyenne de 12 ans datant de la période 2000-2012 sur lesquelles sont rajoutées les variabilités interannuelles. On y observe deux principales zones de creux, la première, s'étendant de fin juin jusque fin juillet, la seconde, plus courte, de fin décembre jusque février. L'indice NDVI peut être corrélé à la quantité de biomasse présente sur la surface d'intérêt. Il se définit à partir de la relation suivante :

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}}$$
 où ρ_{RED} et ρ_{NIR} correspondent aux réflectances des bandes rouge visible et proche infrarouge, respectivement. Il est toutefois nécessaire de préciser que les mesures obtenues par le capteur MERIS sont empreintes d'incertitudes. Cette incertitude est liée à la probabilité d'obtenir une image composite hebdomadaire sans nuages et autres artefacts atmosphériques. Les valeurs réelles du NDVI pourraient donc être modifiées par ces incertitudes, ces dernières étant proportionnelles aux valeurs de l'histogramme en gris (Figure 18). Hormis une légère atténuation des pics et creux du NDVI dans les zones plus humides, peu de variations de ce dernier sont observables entre les zones forestières sèches et les zones forestières marécageuses.

Afin de mieux observer les variations pouvant s'établir entre les saisons sèches et humides, deux images S1 moyennées selon chacune des saisons ont été soustraites l'une à l'autre. Les résultats de cette différence, ventilés selon des écarts-types de ± 1 , $\pm 1,5$ et ± 2 sont reproduits ci-dessous (Figure 19). Bien que quelques patterns linéaires soient identifiables sur la Figure 19 (c), les points résultant de la différence des deux images semblent suivre une distribution spatiale uniforme et aléatoire. Il est dès lors difficile d'identifier des changements de manière significative. Les différences les plus marquées se situent cependant au niveau des cultures agroforestières à proximité des zones habitées.

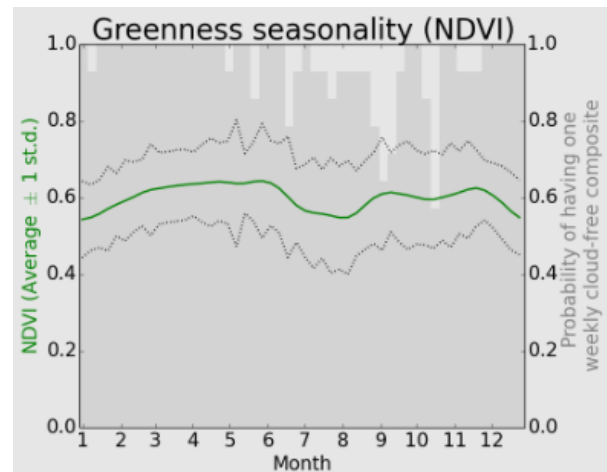


Figure 18 : Évolution du NDVI (vert) et de la probabilité d'acquisition d'une image composite sans nuages (barres grises) sur l'année 2015 - Pokola, Congo. Estimation à partir des données de réflectance acquises par le capteur MERIS (ESA/CCI-LC, UCLouvain, 2015)

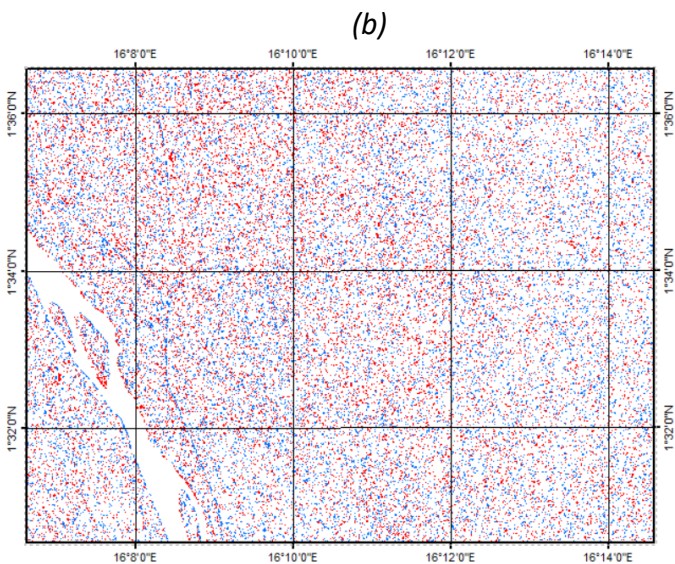
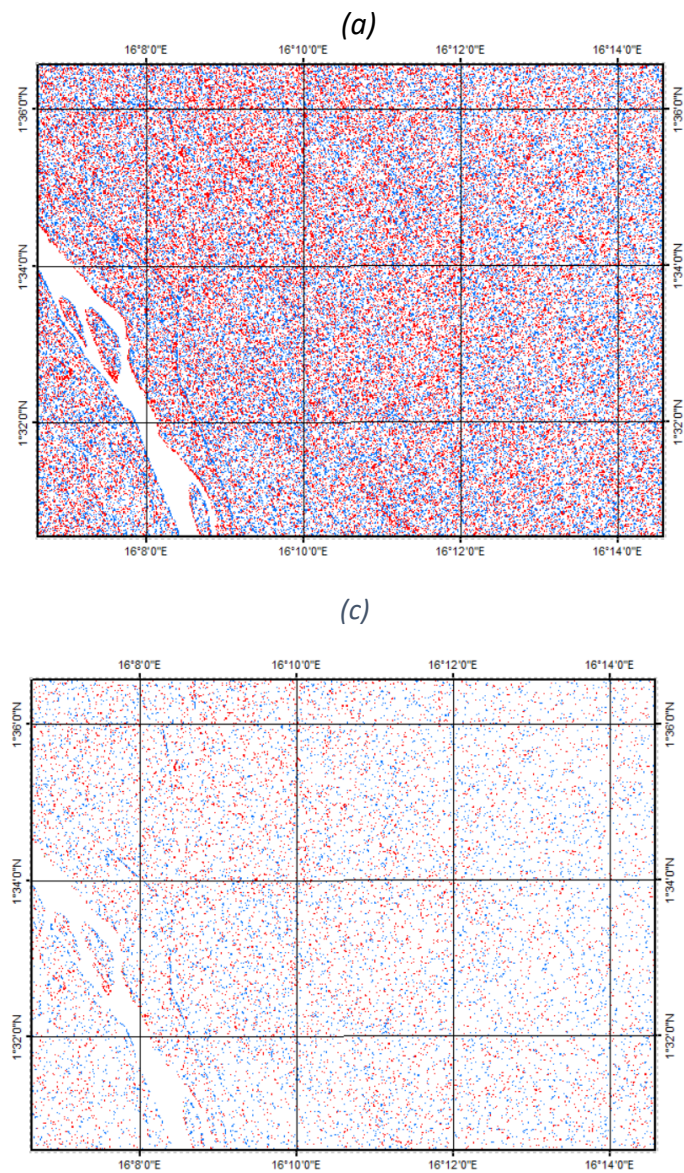
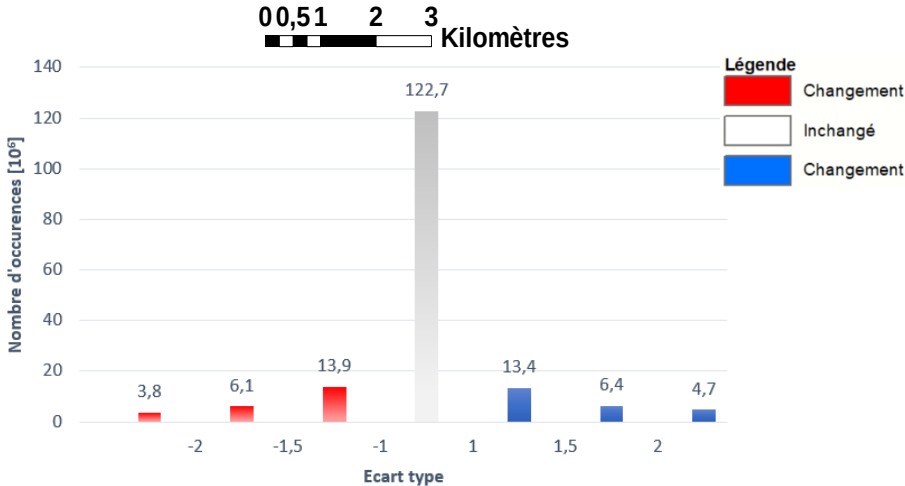


Figure 19 : Comparaison de trois images issues de la différence des coefficients de rétrodiffusion [dB] des deux images S1 moyennées sur les périodes de maximum et de minimum de production de biomasse végétale. Les images résultant de cette différence sont ventilées selon un écart-type de ± 1 (a), $\pm 1,5$ (b) et ± 2 (c).



3.2.1.2 Segmentation

Les images radar ont pour défaut de présenter un ratio signal/bruit relativement faible. En conséquence, les images ont un aspect poivre et sel, un pixel peut présenter un signal très fort tandis que son voisin, a contrario, aura un signal très faible. De telles différences des coefficients de rétrodiffusion sur de courtes échelles spatiales restreignent le potentiel de production d'une classification par pixel de qualité. L'utilisation de filtre permet de limiter cet effet de speckle sans pour autant le supprimer, mais cela a pour conséquence d'également réduire la résolution spatiale de l'image.

Afin de représenter au mieux les différentes classes forestières identifiées sur le terrain et de limiter le bruit inhérent aux images radar, il est nécessaire de segmenter l'image en multiples objets présentant des propriétés homogènes. Le speckle des images radar a également pour effet de restreindre la capacité des algorithmes de segmentation à réaliser un partitionnement de l'image respectant les frontières physiques des éléments composant l'image.

Cette analyse préliminaire cherchera à mettre en évidence l'algorithme présentant les meilleures performances de segmentation. Ces performances seront évaluées sur base de plusieurs paramètres : la facilité d'utilisation de l'algorithme, la capacité à mettre en évidence certains objets d'intérêts, la capacité de l'algorithme à respecter les frontières spatiales et spectrales et le temps de calcul.

3.2.1.2.1 Comparaison de trois techniques de segmentation : SLIC, SLICO, et MRS

Ces techniques sont fort répandues, autant dans l'analyse d'image optique que radar, pour permettre la segmentation d'image (Achanta et al., 2012; Clauss et al., 2018; Clerici et al., 2017; Delgado-Aguilar et al., 2017; Liu et al., 2018; Lohberger et al., 2018).

Simple Linear Iterative Clustering – SLIC

La technique de segmentation SLIC se base sur l'adaptation de la méthode de *clustering* des données en k-moyennes. Elle diffère de cette méthode par deux points cruciaux :

- 1) Les calculs de distance permettant l'optimisation du processus sont réduits en limitant l'espace de recherche à une région proportionnelle à la taille du superpixel (Figure 20).

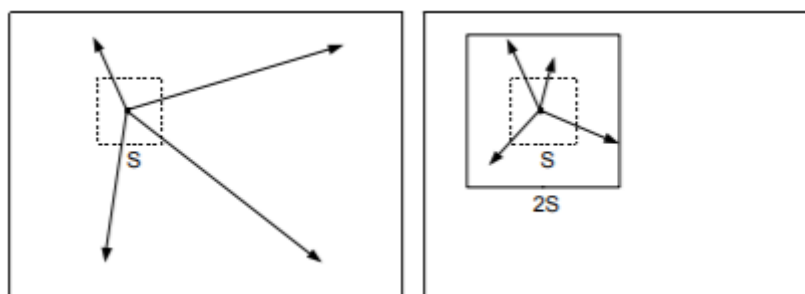


Figure 20 : Ces deux figures représentent l'espace de recherche des deux algorithmes k-moyennes (à gauche) et SLIC (à droite). Dans la méthode conventionnelle k-moyennes, les distances sont calculées depuis chacun des groupements sur tous les pixels de l'image. L'algorithme SLIC limite son espace de calcul des distances du centre de chacun des groupements jusqu'aux pixels compris dans une surface de $2S \times 2S$. Il faut noter que la taille présumée d'un superpixel est de $S \times S$, indiquée par le carré pointillé (Achanta et al., 2012)

- 2) Une mesure de distance pondérée combine les proximités spatiales et spectrales tout en permettant de garder un contrôle sur la taille et la compacité des pixels.

L'approche par superpixel ne dépend que d'un paramètre par défaut : k , le nombre souhaité de superpixels de taille similaire. L'algorithme est initié par la création de k groupements, répartis de manière homogène dans une grille où la distance entre chaque groupement est strictement supérieure à S .

On peut dès lors établir les relations suivantes :

$N = \text{Nombre de pixels dans l'image}$

$k = \text{Nombre de superpixels}$

$\frac{N}{k} = \text{Surface moyenne des superpixels}$

$S = \sqrt{N/k} = \text{Distance entre les centres}$

Chacun des centres des groupements est caractérisé par une position x et y sur l'image ainsi que plusieurs composantes spectrales correspondant à la position définie. Les opérations de partitionnement de l'image suivent ensuite le même processus que l'algorithme k -moyennes.

Deux méthodes de segmentation basées sur cet algorithme ont été testées :

SLIC

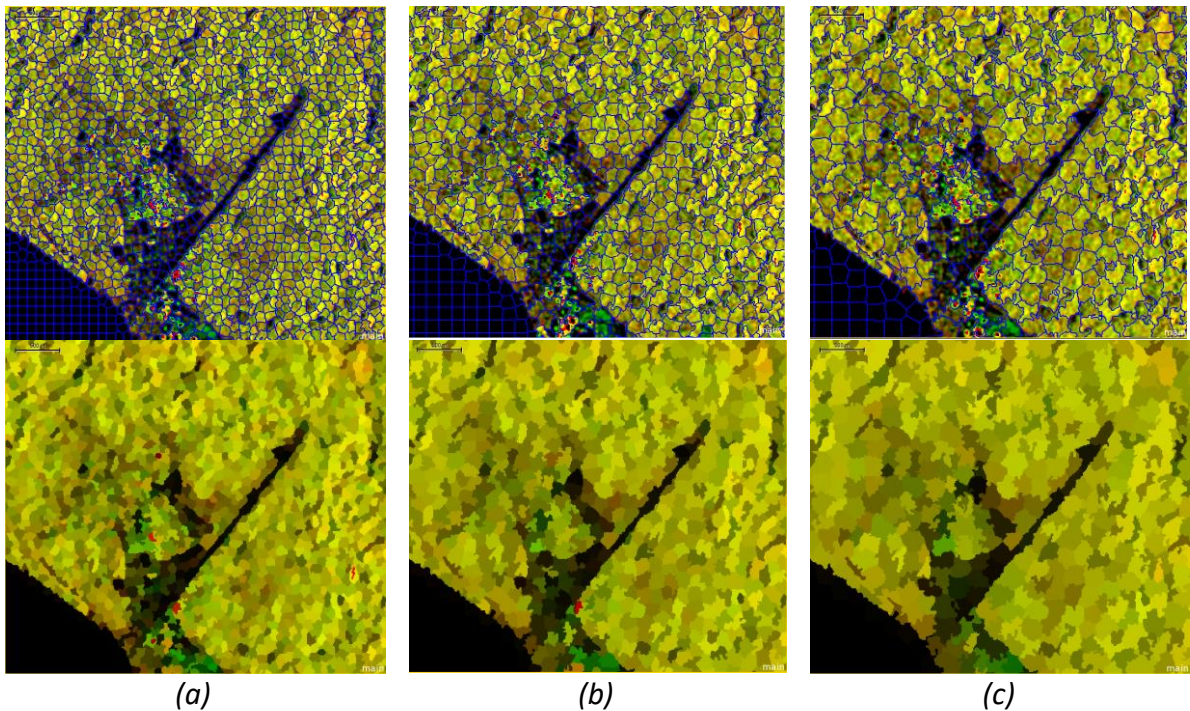


Figure 21 : Comparaison de 3 segmentations réalisées à partir de l'algorithme SLIC selon un facteur de fusion de 25 % et de compacité de 100. Le nombre d'itérations est fixé à 10 pour ne modifier qu'une seule variable, le facteur de forme. Celui-ci est modifié entre les différentes segmentations comme suit : (a) Facteur de 10, (b) Facteur de 15 et (c) Facteur de 20. Les images du dessus correspondent à l'image originale segmentée par superpixels, celle du dessous aux données spectrales moyennées par superpixels. Composition colorée : VV (Rouge), VH (Vert), VV/VH (Bleu)

SLICO

SLICO utilise un paramètre de rugosité variable selon les propriétés texturales de l'image. Ce paramètre permet d'éviter d'avoir des superpixels aux formes très hétérogènes dans les zones fortement contrastées et des superpixels très réguliers dans les zones homogènes.

La principale variable évaluée lors de l'analyse de cette segmentation fut le nombre d'itérations. Au plus élevée est cette variable au plus juste sera la segmentation. En conséquence, différentes segmentations furent réalisées avec 5, 10 et 50 itérations. Hormis le temps de calcul qui s'allongeait avec le nombre d'itérations, les résultats restaient similaires (Figure 22). Il est considéré dans la littérature que 10 itérations suffisent à avoir une segmentation de qualité. Au-delà de dix itérations, il est difficile d'observer des améliorations dans la segmentation (Achanta et al., 2012). Le paramètre de forme fut fixé à sept. Ce dernier indique la taille optimale à atteindre pour la formation d'un superpixel, sept fois sept pixels dans ce cas-ci. Une valeur trop importante de ce paramètre limiterait la capacité de l'algorithme à respecter la variabilité spectrale des différents pixels de l'image.

Le résultat final de cette méthode est la génération d'une segmentation très régulière. Le fait que l'algorithme force la régularité de la taille et de la forme des objets induit régulièrement des erreurs dans les zones fort hétérogènes résultant en une segmentation de qualité insuffisante.

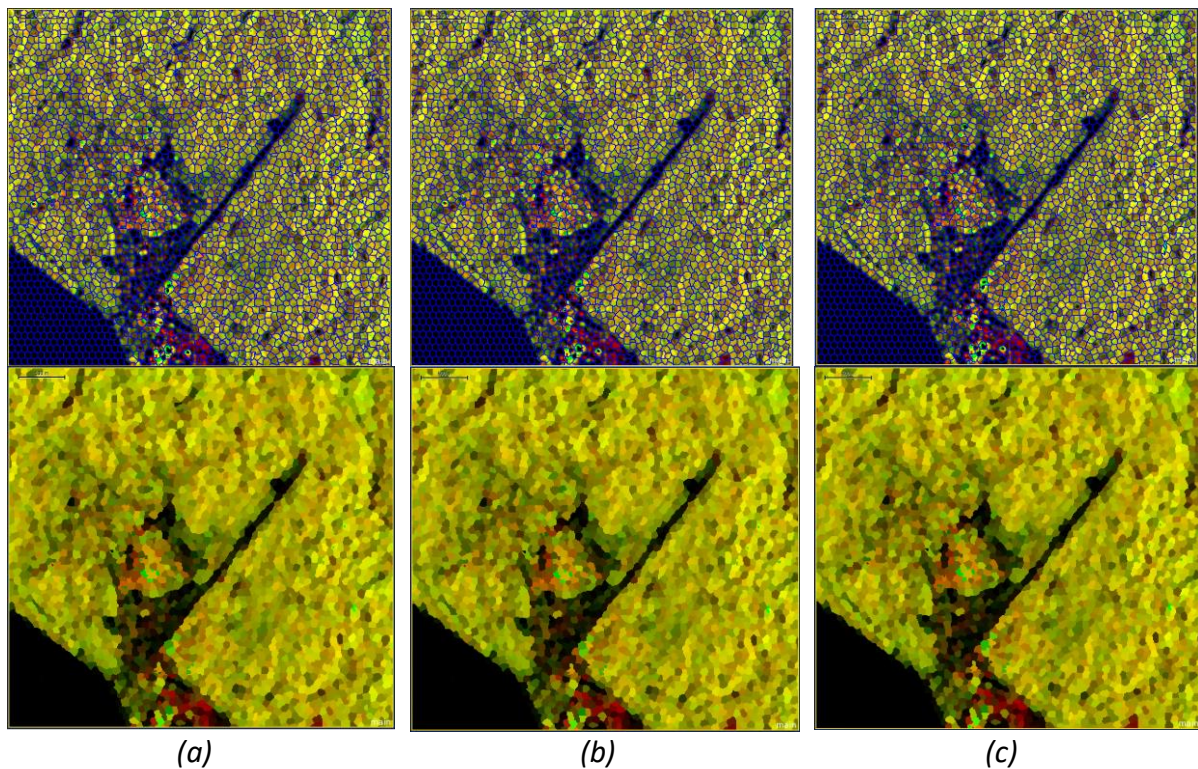


Figure 22 : Comparaison de 3 segmentations réalisées à partir de l'algorithme SLICO selon un facteur de forme de 7 et de fusion de 10 %. Seul le nombre d'itérations est modifié entre les segmentations : (a) 5 itérations, (b) 10 itérations, (c) 50 itérations. Les images du dessus correspondent à l'image originale segmentée par superpixels, celle du dessous aux données spectrales moyennées par superpixels. Composition colorée : VV (Rouge), VH (Vert), VV/VH (Bleu).

Segmentation multirésolution (MRS)

La procédure de segmentation d'image multiéchelle est décrite par Baatz et Schäpe (s.d) comme une technique de fusion de région. Elle s'initie par l'association de paires de pixels formant un objet dans l'image. À chaque étape, une paire de pixels sélectionnés est fusionnée dans un objet plus grand. La décision de fusion est basée sur un critère d'homogénéité local, décrivant la similarité spectrale des objets adjacents.

Le critère d'homogénéité n'est pas un simple critère de fusion ou d'exclusion binaire, il est lié à un « coût de fusion » inhérent à chaque fusion potentielle. Ce coût représente la différence

du degré de similarité de deux objets avant et après la fusion. Pour qu'une fusion soit réalisée, il faut que ce coût soit inférieur à une valeur seuil. Le processus s'arrête finalement dès qu'il n'y a plus de fusion possible. En conséquence, la valeur seuil du coût de fusion – que l'on peut nommer « paramètre d'échelle » – influe directement sur la taille des objets, une plus grande valeur seuil étant plus permissive à la fusion de deux objets résultants en une entité présentant une certaine hétérogénéité.

La segmentation multirésolution est régie par deux processus principaux (Baatz and Schäpe, n.d.) :

- Une décision heuristique déterminant les objets qui seront fusionnés à chacune des étapes.
- La définition d'un degré d'homogénéité des objets à faire correspondre avec le degré d'acceptation du processus de fusion de deux paires de pixels.

Les tests réalisés sur cette technique de segmentation utilisaient un paramètre de compacité de 1, permettant de minimiser le ratio *Périmètre de l'objet / Surface de l'objet*. Le paramètre d'échelle, quant à lui, fut fixé à 30 et permet de limiter la sursegmentation des objets tout en produisant des objets de taille limitée respectant bien les différences spectrales. Le seul réglage sera donc le paramètre de forme, indice de sensibilité de l'algorithme aux valeurs spectrales de chacun des pixels. Ce dernier sera testé avec les valeurs 0.2, 0.4 et 0.6. Au plus la valeur tend vers 1, au moins l'algorithme est sensible aux variations spectrales.

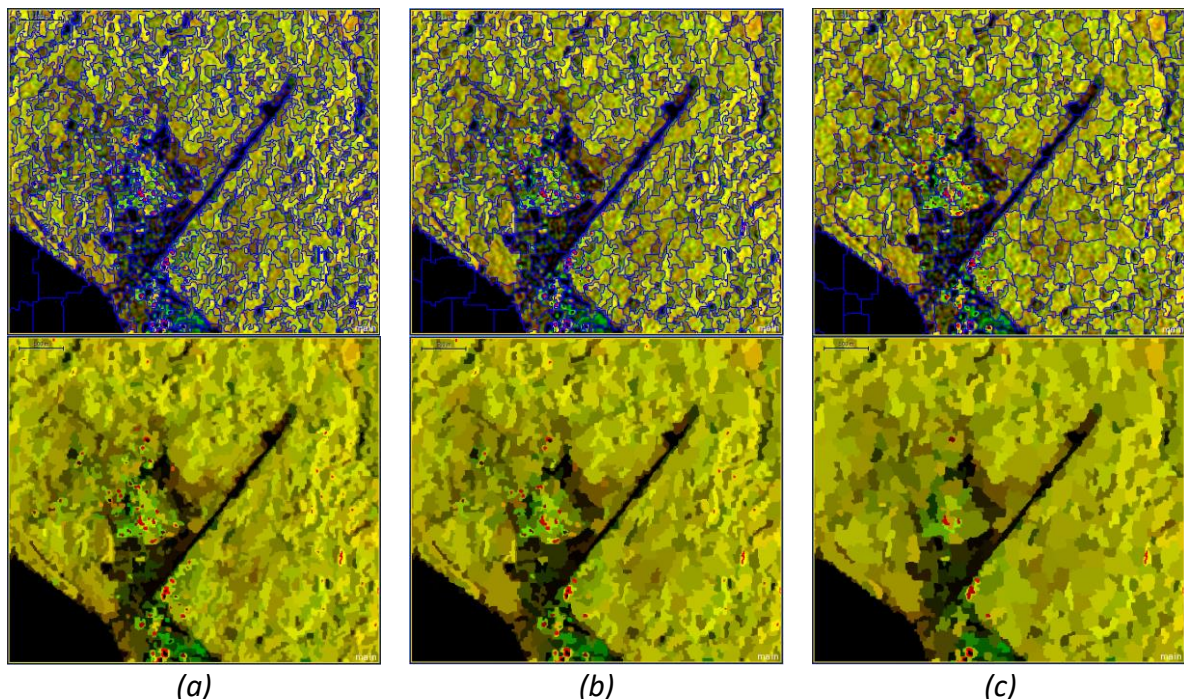


Figure 23 : Comparaison de 3 segmentations réalisées à partir de technique de segmentation multirésolution selon un paramètre d'échelle de 30, un indice de compacité de 1. Seul le paramètre de forme est modifié entre les segmentations : (a) paramètre de forme de 0,2, (b) paramètre de forme de 0,4, (c) paramètre de forme de 0,6. Les images du dessus correspondent aux frontières des objets, délimitées par la segmentation, celles du dessous, aux données spectrales moyennées par objet. Composition colorée : VV (Rouge), VH (Vert), VV (Bleu).

Le résultat final (Figure 23) est caractérisé par la formation d'objets de formes très irrégulières, ceux-ci respectant bien la géométrie des différentes entités présentes dans l'image initiale. Au plus le paramètre de forme est faible, au plus la segmentation génère des objets de forme

irrégulière. Cette technique de segmentation présente le temps de calcul le plus faible (environ deux minutes pour une image de 350 Mb), permettant de traiter rapidement des images de grande dimension. Cette segmentation, bien qu'efficace est soumise à plusieurs contraintes. Premièrement, l'extraction d'un objet d'intérêt est soumise à un set de paramètres bien défini, rendant difficile l'extraction générale d'objets présentant des propriétés différentes. Deuxièmement, selon le facteur d'échelle, il est difficile de trouver le compromis entre une sursegmentation permettant de mettre en évidence les éléments de faible dimension tels que les routes et pistes forestières, mais générant davantage d'objets « parasite » ne correspondant pas à la réalité terrain et une segmentation trop simplifiée ignorant les limites physiques des entités de l'image.

3.2.1.2.2 Synthèse des algorithmes de segmentation

Sont reprises ci-dessous (Tableau 6) les différentes caractéristiques des trois algorithmes de segmentation étudiés : MRS, SLIC et SLICO. Les opérations de segmentation ont été réalisées sur l'ensemble de la zone d'étude soit une surface de 6.600 kilomètres carrés. Les résultats présentés dans les Figure 21, Figure 22 et Figure 23 ciblent l'agglomération de Pokola et ses environs hétérogènes occupés principalement par des cultures d'abattis-brûlis et de jeunes peuplements de parasoliers.

Tableau 6 : Synthèse des performances respectives de la segmentation multirésolution, SLIC et SLICO.

Algorithme	Segmentation multirésolution	SLIC	SLICO
<i>Méthode de base</i>	Coût de fusion	k-moyennes	k-moyennes
<i>Temps de calcul</i>	Rapide, 2 minutes	Fonction du nombre d'itérations 15 minutes – 10 itérations	Fonction du nombre d'itérations 30 minutes – 10 itérations
<i>Nombre de paramètres</i>	3 paramètres : - Echelle - Compacité - Forme	4 paramètres : - Nombre d'itérations - Taille du superpixel - Rugosité - Degré de fusion des éléments de faible dimension	3 paramètres : - Nombre d'itérations - Taille du super pixel - Degré de fusion des éléments de faible dimension
<i>Forme</i>	Hétérogène	Homogène – faible contraste Hétérogène – fort contraste	Homogène
<i>Facilité d'utilisation</i>	Difficile	Moyen	Facile
<i>Performance générale</i>	Efficace, mais nécessite un set de paramètres particuliers pour chaque objet	Efficace	Insuffisante

Des trois algorithmes de segmentation étudiés, l'algorithme MRS présentait les meilleures performances. Son utilisation est cependant complexe, chaque objet d'intérêt devant être extrait par des paramètres de segmentation différents. Afin de répondre à cette difficulté d'extraction des objets d'intérêts, Drăguț, et al. (2010) développèrent un nouvel algorithme de segmentation basé sur MRS permettant d'identifier le paramètre d'échelle optimal à la segmentation d'une image.

3.2.1.2.3 Estimation du paramètre d'échelle (ESP)

Cette méthode, basée sur les travaux de Drăguț, et al. (2010), de Drăguț et Eisank (2012) et développée par Drăguț et al. (2014) permet la segmentation d'image via l'algorithme de segmentation multirésolution. Le principal défaut de cette dernière méthode de segmentation réside dans la difficulté de trouver le paramètre d'échelle optimal permettant une segmentation de qualité. Drăguț explique que l'algorithme ESP considère le paramètre d'échelle comme étant une variable spatiale intrinsèquement liée à la structure de l'image. Il permet dès lors de s'affranchir du choix de celui-ci, l'algorithme l'évaluant et l'ajustant de lui-même.

L'algorithme de segmentation ESP repose sur l'exploitation d'un paramètre de variance locale (LV) capable de détecter les transitions d'échelle dans les données géospatiales. L'outil détecte le nombre de couches ajoutées à un projet et les segmente de manière itérative selon une approche ascendante ou descendante où le facteur d'échelle de la segmentation augmente selon un incrément constant renseigné par l'utilisateur. La valeur du paramètre de variance locale est calculée à chacune des itérations et sert de condition pour mettre fin au processus de segmentation : lorsque la valeur LV d'un niveau concerné est égale ou inférieure à la valeur de la précédente itération, l'itération se termine et les objets du niveau concerné sont segmentés (Drăguț et al., 2014).

L'algorithme ESP est défini par les paramètres suivants :

Tableau 7 : Paramètres conditionnant l'algorithme ESP.

Paramètre	Fonction
<i>Hiérarchie</i>	Définit la manière avec laquelle doit être généré le paramètre d'échelle. Celui-ci peut suivre une hiérarchie à plusieurs niveaux indiquée par des paliers ou être généré de manière indépendante, basée seulement sur la valeur du pixel.
<i>Type de hiérarchie</i>	Ascendante (Bottom-up) : Utilise différents critères basés sur l'image pour définir des groupes de pixels cohérents ayant une forte probabilité d'être associés ensemble. Descendante (Top-down) : Utilise la représentation d'objets donnés en tant qu'exemples pour identifier des objets particuliers dans l'image d'intérêt.
<i>Incrément par niveau</i>	L'algorithme de base sélectionne trois niveaux définis par un palier (paramètre d'échelle initial) sur lesquels seront réalisées les itérations. Ce paramètre permet de modifier la taille de l'incrément, par niveau.
<i>SP initial par niveau</i>	Valeur du paramètre d'échelle initial, définissant les paliers des niveaux.
<i>Compacité</i>	Définit le ratio Périmètre/Surface des segments.
<i>Forme</i>	Définit la sensibilité de l'algorithme à la composante spectrale des pixels.
<i>Nombre de boucles</i>	Définit le nombre d'itérations devant être réalisées.

Le processus de segmentation se clôture par la génération d'un graphe (Figure 24) décrivant l'évolution de LV et du taux de changements pour chacune des itérations du paramètre d'échelle. Bien que la Figure 24 n'ait pas pu mettre en évidence un paramètre d'échelle optimal pour la segmentation de l'image, la segmentation résultante fut conservée, car elle permit la mise en évidence de certains objets d'intérêt tels que les zones agricoles et forestières anthropisées (Figure 25).

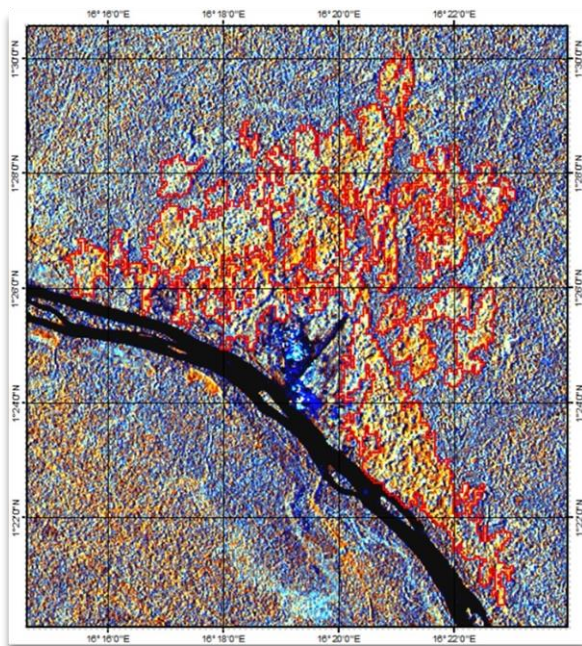
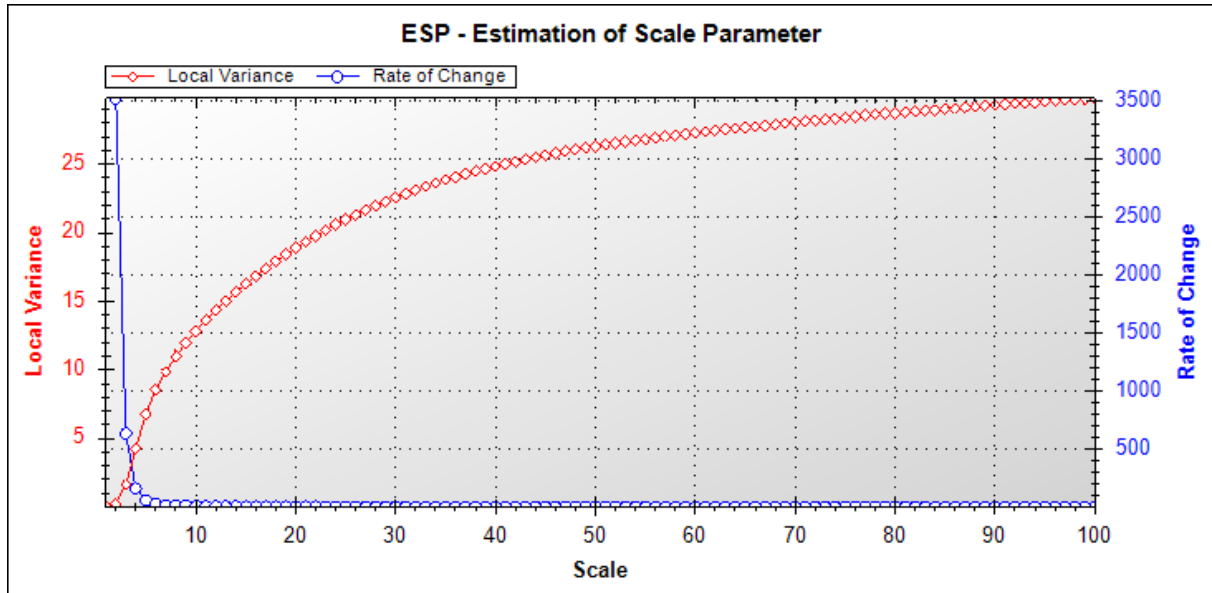


Figure 24 (au-dessus) : Évolution de la variance locale en fonction du paramètre d'échelle selon la méthode développée par Drăguț et al. (2010) et améliorée par Drăguț et al. (2014). Cette évaluation de LV est réalisée sur base de l'ensemble de la segmentation de la zone d'étude.

Figure 25 (à gauche) : Résultat de la segmentation par ESP2. Les zones anthropisées aux alentours de Pokola (Rouge) ont été extraites de la segmentation à partir d'une classification supervisée sur la valeur moyenne du coefficient de rétrodiffusion de la bande VV, VH et VV/VH ainsi que sur les paramètres texturaux du contraste, de la corrélation et des moyennes GLCM. Echelle : 1 : 75 000.

Cette segmentation fut réalisée à partir des paramètres suivants : Hiérarchisée, ascendante, paramètre d'échelle initial : 1, pas de l'itération : 1, coefficient de forme : 0,1, compacité : 1, nombres d'itérations : 100.

La méthode de segmentation MRS-ESP2, malgré un temps de calcul très important (72 heures pour 100 itérations de segmentation d'une image de 350 Mb) présente des performances de segmentation similaire à celle de MRS. MRS-ESP2 permet toutefois la mise en évidence d'objets difficilement extractibles par MRS. En considération de leurs performances respectives, ces deux algorithmes seront utilisés pour réaliser l'ensemble des opérations de segmentation.

3.2.1.3 Analyse de la rugosité et des textures

Afin de faciliter la discrimination des forêts primaires et des forêts exploitées, un accent fut mis sur l'analyse de la rugosité ainsi que l'identification des couches texturales les plus pertinentes pour mener à bien cette discrimination. La Figure 26 renseigne sur la distribution des données de rugosité selon les différentes zones visitées. En faisant correspondre ces données sur la rugosité de la canopée avec l'année d'exploitation (Carte 5, Figure 26), on observe une légère tendance à l'augmentation de cette dernière dans les milieux qui ont été exploités récemment, conformément à la théorie (Figure 26). Le jeu de données récolté présente toutefois de nombreuses lacunes. En effet, aucune donnée sur la rugosité n'a été récoltée pour les peuplements dont l'exploitation remonte aux années 1992 à 2014. Cette absence de données s'explique notamment par les efforts qu'entretient la CIB pour limiter les opérations de braconnages perpétrées par les populations locales. Dès que l'exploitation d'une parcelle arrive à son terme, les pistes d'exploitations et autres infrastructures routières menant à la parcelle en question, sont fermées. Cela permet de limiter la mobilité des braconniers et ainsi, de protéger les nombreuses espèces animales convoitées par ces derniers. En conséquence, le réseau routier ne dessert que les milieux forts fréquentés ou les parcelles forestières sujettes aux récentes et futures exploitations. Dès lors, aucune donnée n'a pu être collectée sur les peuplements dont l'exploitation remonte aux années 2000 – 2013.

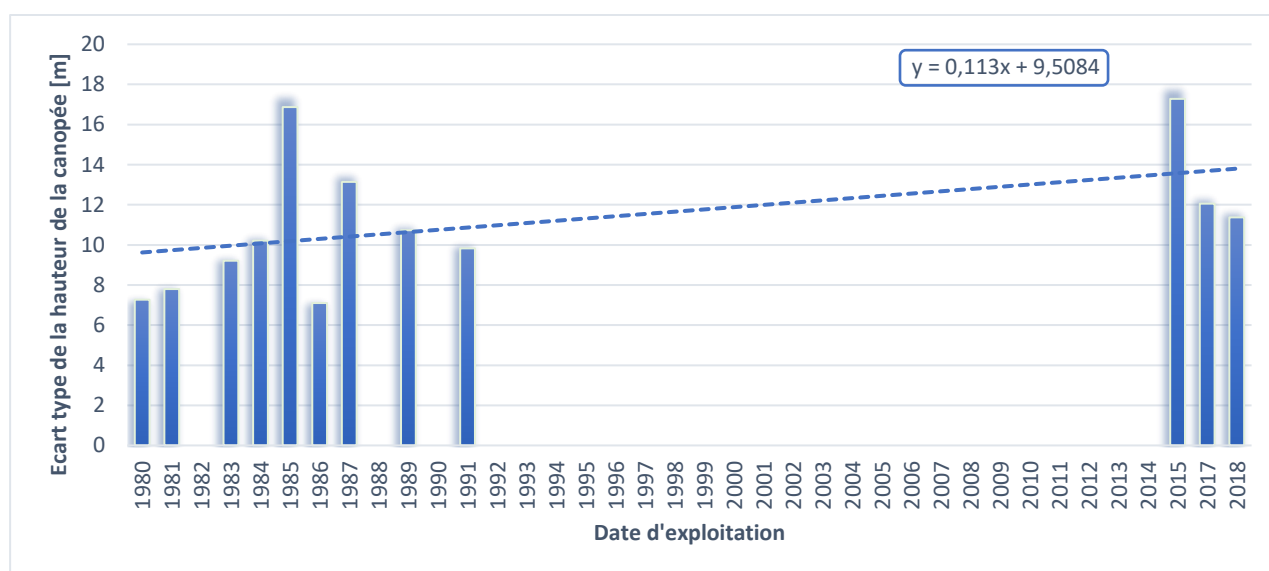
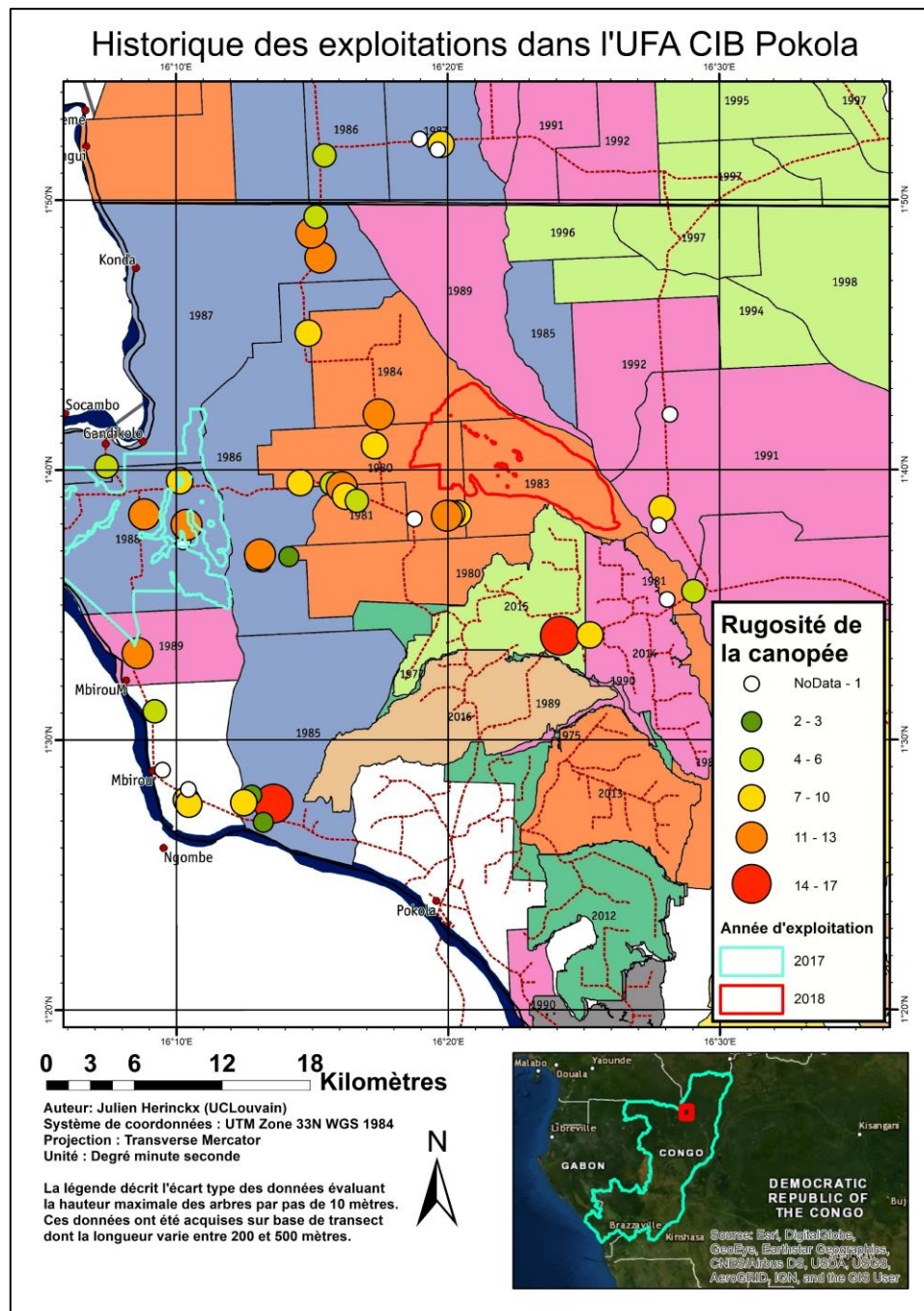


Figure 26 : Moyenne des écarts-types de la hauteur maximale de la canopée distribuées selon les dates d'exploitations des parcelles échantillonnées

Afin de pouvoir identifier la variable permettant de séparer les écosystèmes forestiers récemment exploités de ceux anciennement exploités, une analyse des variables texturales fut considérée. Les données considérées ont été acquises sur base d'un tampon de 50 mètres appliqué sur chacun des transects établis dans les zones d'intérêts. Les données collectées ont ensuite été distribuées en deux groupes, les peuplements ayant connu une exploitation récente (2015-2018) et les peuplements dont l'exploitation remonte à plus de 20 ans (1980-1991). L'extraction des moyennes et écarts-types de chacun des deux groupes permet la réalisation d'un test de Student visant à confirmer ou à infirmer l'hypothèse que les moyennes des deux groupes sont différentes et, dès lors, que l'exploitation a un impact sur le signal apporté par les données texturales (Tableau 8).



Carte 5 : Distribution des mesures de rugosité de la canopée pour chacune des zones visitées. Les dates mentionnées en arrière-plan de la figure correspondent aux années d'exploitation des parcelles.

Tableau 8 : Résultat de comparaison des moyennes de deux jeux de données, l'un fixé sur les forêts anciennement exploitées, l'autre fixée sur les forêts récemment exploitées évalués selon le 95^{ème} et le 99^{ème} percentile.

$H_0 : \mu_1 = \mu_2$	Texture				
α	Contraste	Dissimilarité	Homogénéité	ASM	Maximum de probabilité
0,05	Rejet	Rejet	Rejet	Rejet	Rejet
0,01	Acceptation	Rejet	Rejet	Acceptation	Acceptation
α	Energie	Entropie	GLCM Mean	GLCM Std	GLCM Corrélation
0,05	Rejet	Acceptation	Acceptation	Rejet	Acceptation
0,01	Rejet	Acceptation	Acceptation	Acceptation	Acceptation

En considérant l'hypothèse nulle que les deux jeux de données n'ont pas de différence significative, on observe que pour une évaluation réalisée au 95^{ème} percentile, les informations texturales liées à l'entropie, aux moyennes GLCM et à la corrélation sont les seules à ne pas pouvoir discriminer les deux types forestiers. Pour une évaluation réalisée au 99^{ème} percentile, seules les informations concernant la dissimilarité, l'homogénéité et l'énergie permettrait de discriminer les deux types forestiers. De manière à sélectionner le set de textures le plus pertinent, une analyse de la corrélation de chacune des textures fut réalisée. La Figure 27 démontre que ces trois données sont fortement corrélées l'une à l'autre. En conséquence, le set de textures liées à la dissimilarité, aux moyennes GLCM et à l'écart-type GLCM, dont la corrélation est faible entre chacune des couches fut sélectionné pour servir d'input aux classifications. Ce choix de textures résulta en la composition colorée suivante (Carte 6), cette dernière permettant de mettre en évidence les peuplements de limballis.

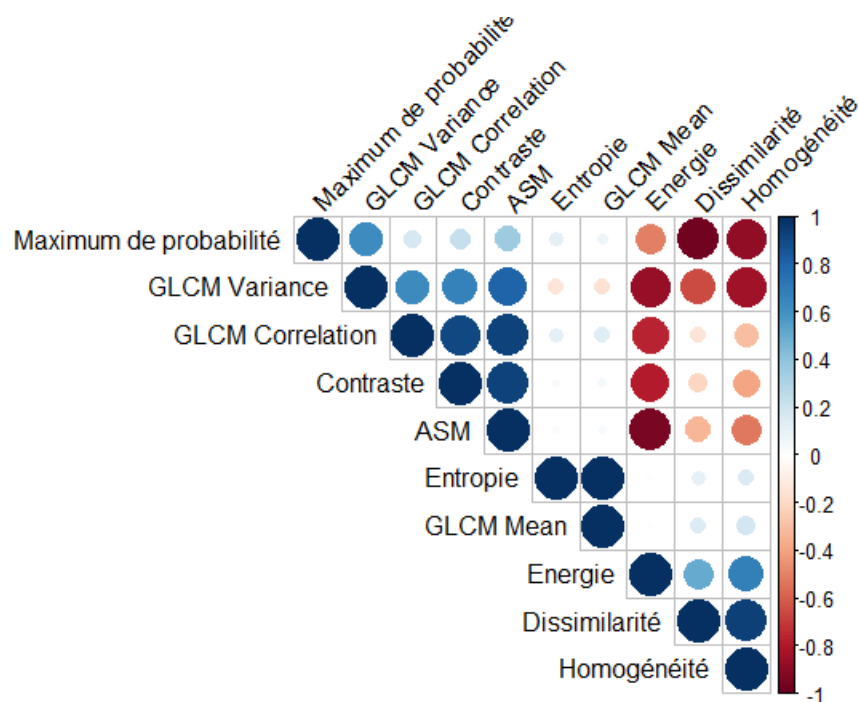
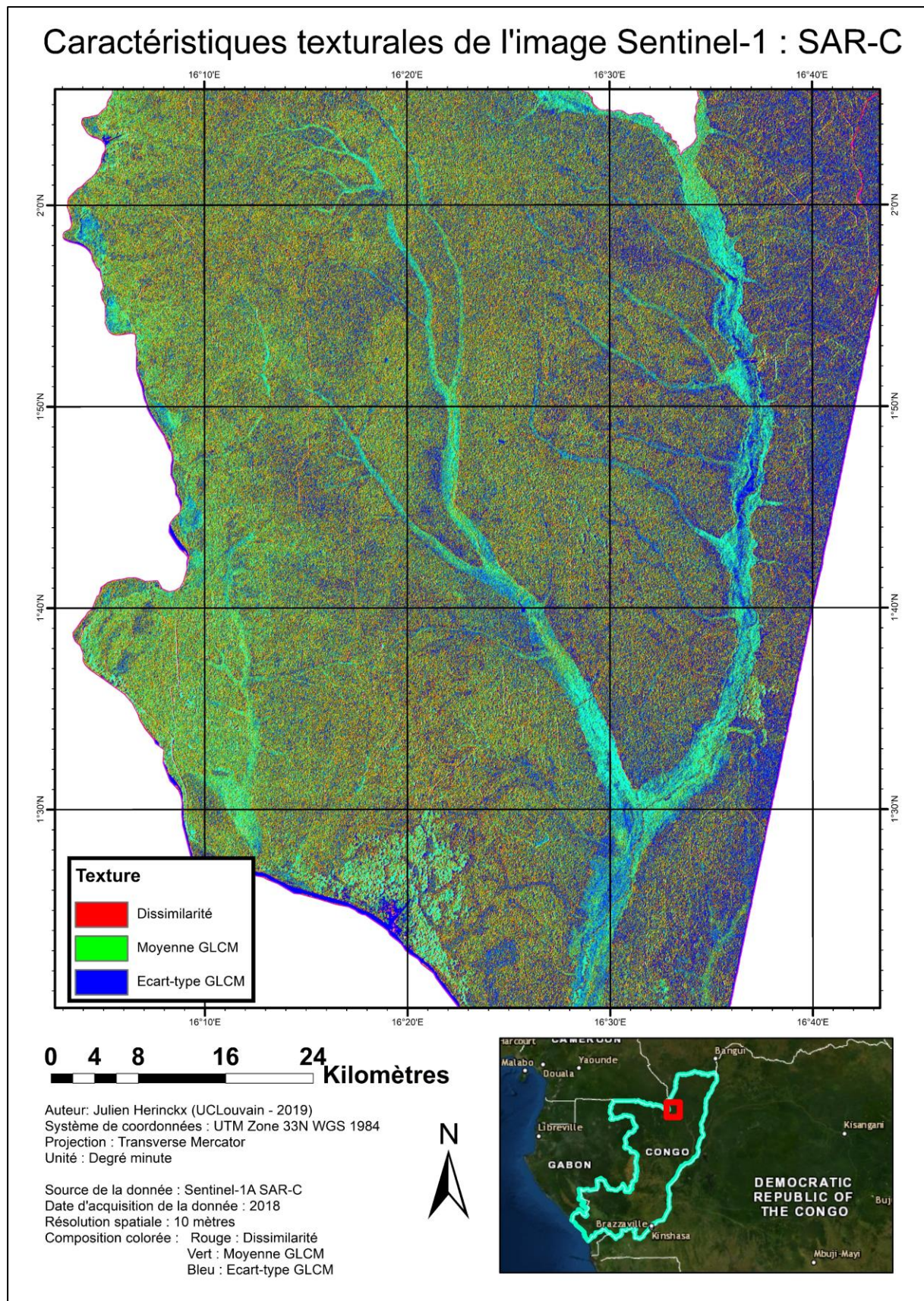


Figure 27 : Matrice de corrélation des différentes caractéristiques texturales de l'image.



Carte 6 : Composition colorée des trois textures permettant de faciliter l'identification des peuplements de limballis.

3.2.2 Traitements

En considération des différentes analyses préliminaires réalisées à la section précédente, la méthodologie appliquée se décrira comme indiqué par l'organigramme suivant (Figure 28). Les données de base (SRTM, S1) une fois rééchantillonnées selon les bonnes résolutions et après applications des différents filtres speckle, serviront, d'une part, d'input pour la réalisation d'une classification par pixel suivie d'opérations post-classification et d'autre part, d'input pour la réalisation d'une classification par objet avec au préalable une segmentation de ces données. Les performances respectives de chacune des classifications seront ensuite mises en évidence par les opérations de validation de ces dernières. Les différentes opérations de traitement indiquées dans l'organigramme sont approfondies dans la suite du document.

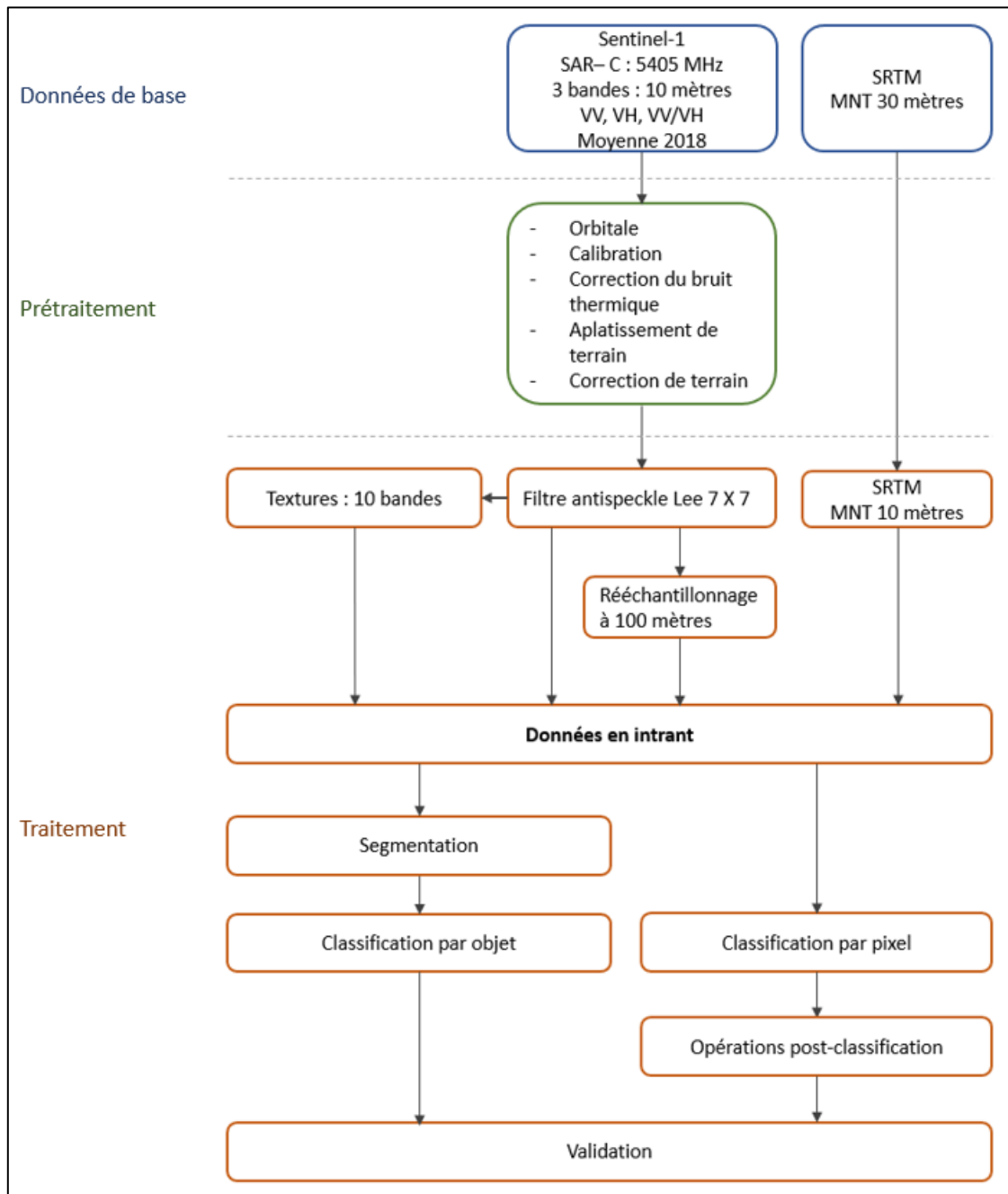


Figure 28 : Organigramme des traitements effectués sur base des données SAR-C

Filtre speckle

Des multiples filtres testés sur les images SAR-C et SAR-L, le filtre Lee présentait le meilleur résultat. Il permet de filtrer les effets de speckle de manière qualitative et ce, en ayant un impact modéré sur la résolution spatiale. Développé par Jong Sen Lee en 1981, ce filtre utilise les statistiques locales pour préserver les détails de l'image. Le filtre de Lee fonctionne de manière adaptative sur base de la variance : si la variance de la surface est faible alors le filtre exécute une opération de lissage de la surface. Ce lissage n'est pas exécuté dans le cas de variance importante. Cette propriété adaptative permet de conserver des détails dans les zones présentant un contraste faible tout comme dans les zones à contraste fort. Le principal défaut de ce filtre est qu'il ne permet pas la suppression efficace du speckle à proximité des frontières d'objets (Shastri, 2015). Le modèle mathématique du filtre de Lee est donné par les équations suivantes :

$$Img(i,j) = Im + W * (Cp - Im) \quad W = \sigma^2 / (\sigma^2 + \rho^2)$$

Avec :

<i>Img</i>	Valeur du pixel après passage du filtre	<i>W</i>	Fenêtre filtrante
<i>Im</i>	Intensité moyenne de la fenêtre filtrante	σ	Variance du pixel
<i>Cp</i>	Valeur du pixel central	ρ	Variance du bruit additif

Moyenne

Les images SAR-C ont été moyennées sur les périodes de saison sèche. Bien qu'il n'existe que de faibles différences entre les images de saison sèche et de saison humide, l'utilisation d'images moyennées sur une saison permet de limiter l'impact des variations hydrologiques. Le choix s'est porté sur la saison sèche afin que les zones marécageuses soient à leurs étiages. Cela permet de délimiter au mieux les zones inondées de manière permanente.

Conversion d'unité

Les données d'amplitude Sentinel-1 sont initialement distribuées sur une échelle de faible dimension, comprise entre 0 et 0,27 pour la bande VV ; 0 et 0,06 pour la bande VH. Afin de faciliter leur manipulation, les données initiales (x_i) sont converties selon deux autres échelles, la première en décibel en y appliquant la formule suivante :

$$x_{dB} = 10 \log_{10}(x_i)$$

La seconde en 8-bit, permettant d'avoir des données distribuées sur 256 valeurs différentes. Cette dernière engendre une perte d'information qui fut considérée comme négligeable pour le traitement des données. Les données PALSAR téléchargeables sont déjà distribuées sur une échelle 8-bit.

3.2.2.1 Stratification de la segmentation

Les différentes données servant d'input pour la classification par objet nécessitent d'être segmentées. L'extraction d'objets étant réalisée sur base de deux algorithmes (MRS et MRS-ESP2), les opérations de segmentation ont été effectuées de façon stratifiée. Ces opérations se basent selon le schéma de la Figure 29 et se décrivent comme suit :

- 1) Création d'un masque « marécage et forêt marécageuse » sur base des données SAR-C et des trois textures apportant le plus d'informations agrégées à 100 mètres de résolution. De manière à faciliter les opérations de segmentation, un filtre passe-bas par fenêtre mobile 5X5 fut appliqué. La polarisation VH fut considérée comme la plus efficace pour mettre en évidence les milieux présentant un régime hydrique particulier. Les marécages ont finalement été extraits de la segmentation en sélectionnant tous les objets dépassant un seuil de 160 au paramètre textural de la variance GLCM, distribuée sur une échelle 8-bit.
- 2) Une segmentation plus fine fut réalisée sur le masque des forêts marécageuses en utilisant l'algorithme de segmentation multi-résolution. Les différents écosystèmes composant les forêts marécageuses ont ensuite été classés via l'algorithme *Random Forest*. Les paramètres de segmentation sont renseignés à la Figure 31.
- 3) Les zones anthropisées ont été extraites grâce à l'algorithme de segmentation MRS-ESP2, une classification supervisée par maximum de probabilité des objets résultants a permis la séparation de ces derniers en trois classes distinctes.
- 4) Le fleuve fut extrait via une classification supervisée par maximum de vraisemblance.
- 5) Les différents types forestiers furent segmentés sur base de l'algorithme de segmentation multi-résolution et uniquement à partir des données texturales de dissimilarité, des moyennes et variances GLCM, cette composition permettant la mise en évidence des forêts de limballis. Les paramètres de segmentation sont renseignés à la Figure 31.
- 6) Les données PALSAR ont permis l'extraction des eyangas, des zones à sols hydromorphes inondées de manière permanente ou semi-permanente et dont le seul apport d'eau est l'eau de pluie. L'extraction fut réalisée sur base d'une classification supervisée réalisée par l'algorithme *Random Forest*. La classification résulta en un ensemble d'entités qui ont ensuite été polygonisés, les objets de dimension inférieure à 3 pixels (1875 m²) et supérieure à 150 pixels (90 000 m²) furent supprimés. Le seuil inférieur permet de supprimer les effets parasites de la classification tandis que le seuil supérieur correspond aux dimensions du plus grand eyanga de la région. Ce dernier seuil permet la suppression des zones marécageuses, dont la taille est beaucoup plus importante.

Une fois ces différentes opérations de segmentation réalisées, l'ensemble des objets est recombinaison selon la hiérarchie indiquée par la Figure 30. Les classes situées au sommet de la figure se superposent aux classes situées au bas de la figure.

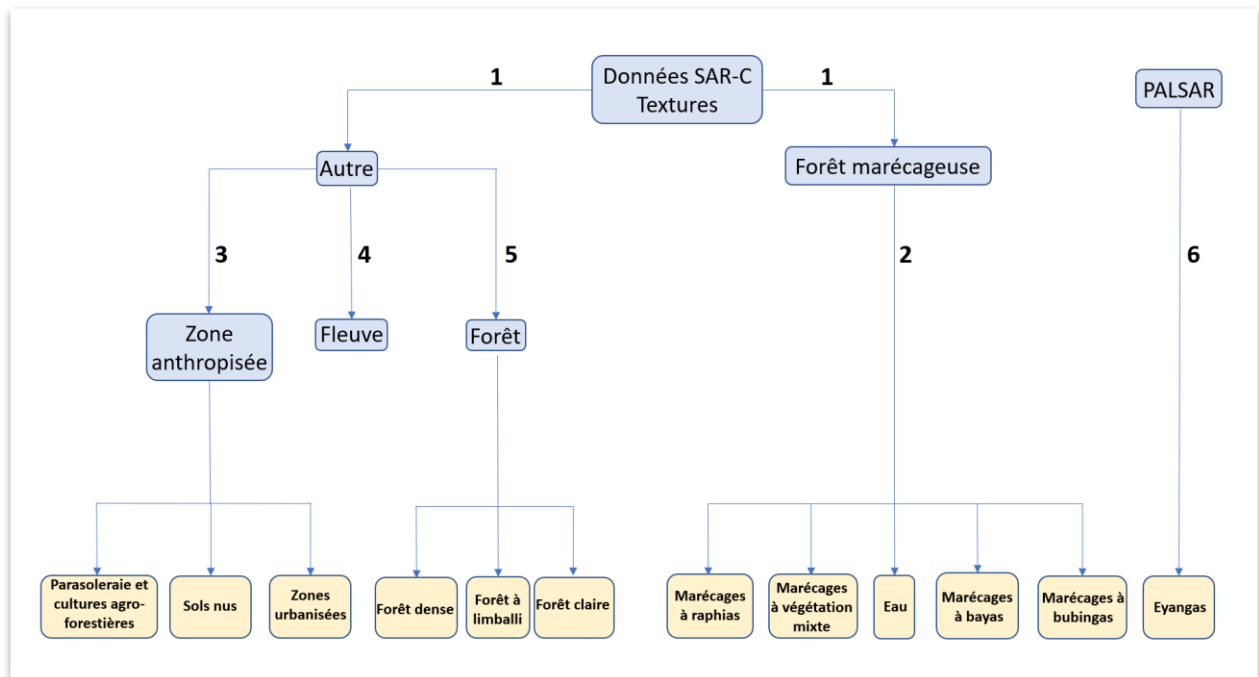


Figure 29 : Schéma suivi pour la réalisation de la segmentation stratifiée

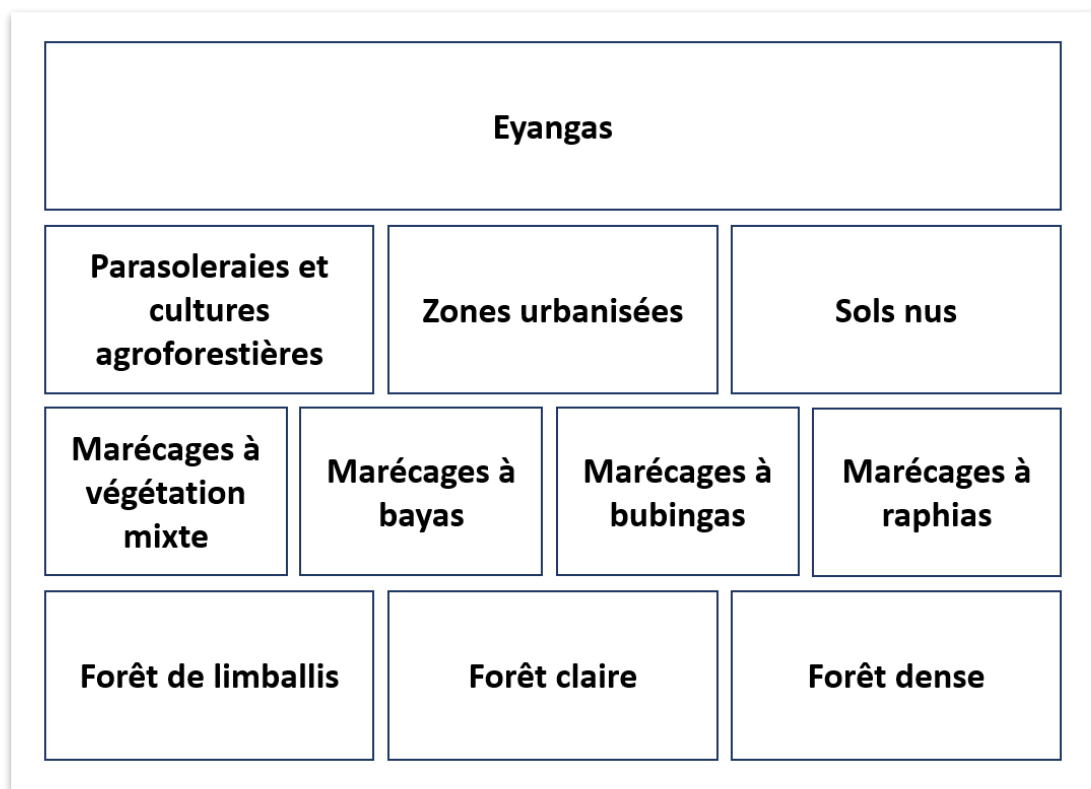


Figure 30 : Stratification des différentes classes résultant de la segmentation. Les classes situées au sommet de la figure se superposent aux classes situées en bas de la figure

Segmentation MRS	Paramètre d'échelle	Paramètre de forme	Indice de compacité
Forêt marécageuse	100	25	1
Forêt sèche	100	8	1

Figure 31 : Paramètres des segmentations fines des forêts sèches et marécageuses

3.2.2.2 Classification

Deux méthodes de classifications ont été suivies, la première est réalisée par pixel tandis que la seconde, réalisée par objet, exploite les objets issus de la segmentation. Les deux classifications ont été réalisées à partir de l'algorithme de classification *Random Forest*. Chacune des aires servant pour l'entraînement de l'algorithme de classification sont déterminées par interprétation visuelle des données S1, en observant les signatures spectrales de chacune des classes identifiées lors de la campagne de terrain.

Différents inputs ont été utilisés pour mener à bien ces opérations de classification :

Tableau 9 : Jeux de données utilisées pour les différents processus de classification.

Matériel	Nombres de bandes	Donnée		
Sentinel-1 : SAR-C	3 Bandes	Polarisation VV, VH et VV/VH		
Textures : SAR-C	10 Bandes	Contraste	Homogénéité	Energie
		Dissimilarité	Max probabilité	ASM
		Entropie	GLCM Mean	GLCM STD
		GLCM Corrélacion		
Shuttle Radar Topography Mission (SRTM)	1 Bande	Données altimétriques fournies par la NASA et la National Geospatial-Intelligence Agency (NGA)		

Classification par objet

Les opérations de classification par objet sont réalisées de manière hiérarchique et stratifiée. Les premières segmentations servaient à extraire les objets principaux de l'image tels que les marécages ou les zones anthropisées. Sur ces extractions ont ensuite été réalisées des segmentations plus fines permettant de mettre en évidence les différences plus subtiles au sein de chacun des objets. Les textures des données SAR-C ont ensuite été extraites à partir de chacun des objets issus de la segmentation, et ce, sous forme de moyenne par objet. Afin de faciliter l'identification des zones marécageuses présentant une topographie particulière, les données SRTM furent ajoutées au set de données permettant la classification.

Classification par pixel et opérations post-classification

Les images radars ont pour propriété d'avoir une apparence dite « poivre et sel ». Deux pixels juxtaposés peuvent présenter un signal très différent l'un de l'autre alors qu'ils correspondent à une seule et même entité physique. Afin de limiter cet effet, la résolution spatiale de l'image fut dégradée à 100 mètres, soit un dixième de la résolution initiale. Bien qu'une telle opération permette de limiter cet effet poivre et sel, elle reste cependant insuffisante pour obtenir une image homogène sur de courtes distances spatiales. En conséquence, les classifications résultant du traitement de ces images peuvent présenter un aspect très granuleux, pouvant mener à des erreurs d'interprétation. Afin d'observer l'importance de chacune des couches mises en input, quatre classifications seront considérées. La première se basera essentiellement sur l'utilisation des données SAR-C tandis que la seconde combinera les données SAR-C avec les données SRTM. La troisième et quatrième incorporeront les données texturales en utilisant la totalité des bandes pour la troisième et seulement les bandes dissimilarité, de moyenne GLCM et de variance GLCM pour la quatrième.

Il existe plusieurs méthodes post-classification permettant de limiter les effets granuleux des classifications par pixels. Bien que ces méthodes permettent de simplifier la lecture de la classification et de corriger les erreurs de cette dernière, elles peuvent également être responsables d'une perte d'information. La combinaison de ces méthodes post-classification

avec une dégradation de la résolution de l'image pourrait avoir un impact non négligeable sur l'exactitude des informations transmises par la classification. Il faut donc manipuler ces dernières avec précaution.

Filtre à majorité

Ce filtre permet le nettoyage des classifications présentant trop de bruit par le passage d'une fenêtre mobile. Il observe un pixel et attribue à ce pixel la valeur majoritairement présente des 4 ou 8 pixels voisins. Les pixels isolés sont dès lors convertis à la classe majoritairement présente dans la fenêtre. Les résultats théoriques d'une telle opération sont représentés par la Figure 32.

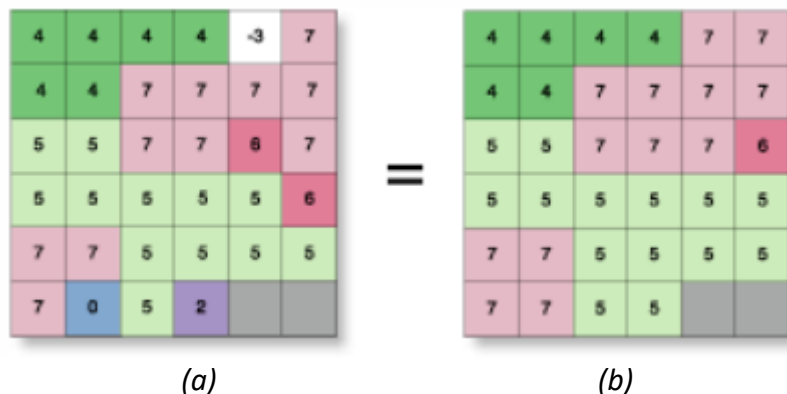


Figure 32 : Visualisation de l'impact du passage d'un filtre à majorité sur une image en entrée (a) et son résultat après application de l'outil (b) par analyse des 4 pixels avoisinants le pixel d'intérêt. Figure issue du panneau d'aide des fonctions de ArcMap10.3.

Nettoyage de limite

Cet outil analyse sur des fenêtres de taille limitée la classe de pixels majoritairement présents. Une fois la classe identifiée, l'outil passe par deux étapes de dilatation et d'érosion des frontières de la classe la plus abondante. La frontière de cette classe est donc étendue à une portée d'un pixel supplémentaire. L'outil réduit ensuite celle-ci en sélectionnant les pixels qui étaient entièrement entourés par des pixels de valeur identique lors de la phase d'expansion. Une fois cette opération terminée, l'outil passe à la seconde classe la plus représentée au sein de la fenêtre de calcul et le cycle recommence. Les résultats théoriques de cette opération sont représentés à la Figure 33.

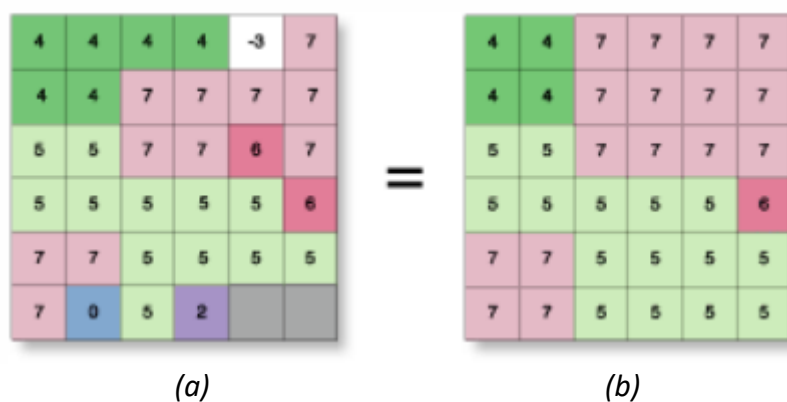


Figure 33 : Visualisation de l'impact de l'utilisation de l'outil "nettoyage des limites" sur une image en entrée (a) et son résultat après application de l'outil (b). Figure issue du panneau d'aide des fonctions de ArcMap10.3.

3.2.2.3 Validation

La validation des classifications par pixel fut réalisée en premier lieu à partir des données récoltées lors de la campagne de terrain. Des cinquante-deux zones initialement analysées, seules 49 d'entre elles furent utilisées pour la validation, les trois autres zones étant trop hétérogènes et peu représentatives d'une classe particulière. Les 49 zones ne couvrant pas l'entièreté des classes présentes dans la classification, d'autres zones de validation ont été ajoutées. Ces dernières concernent principalement les milieux anthropisés, urbanisés ainsi que les zones de sols nus. Aucune analyse terrain ne fut réalisée sur ces milieux. Ce jeu de données étant insuffisant à la représentation de chacune des classes, 100 points furent distribués par classe de manière aléatoire stratifiée et égalisée. Une telle distribution permet d'avoir un nombre égal de points par classe représentée dans la classification. Afin d'éviter les autocorrélations spatiales pouvant modifier artificiellement les résultats de la classification, les points furent séparés d'une distance minimale de 100 mètres. Une telle imposition a également pour conséquence de diminuer le nombre de points des classes couvrant une faible surface et ne pouvant pas accueillir les 100 points initialement prévus.

La classification par objet fut également validée à partir des données de terrain. Aucune donnée n'ayant été acquise sur les milieux anthropisés, urbanisés et les zones de sols nus, quinze points de validation ont été rajoutés sur chacun de ces milieux. Afin d'avoir un jeu de données également représentatif de chacune des classes, un troisième set de données fut utilisé pour parvenir à la création d'une matrice de confusion pertinente. Ce set de données supplémentaire fut équilibré de manière à avoir un nombre égal de points de validation par classe.

Ce protocole de validation permet la création de matrices de confusion sur lesquelles sont évalués différents paramètres permettant de caractériser la qualité des classifications.

F-score

Le F-score ou F1 score est une moyenne harmonique de l'exactitude et du *recall* des différentes classes, il témoigne de la qualité d'une classification et se définit par la formule suivante :

$$F - score = 2 * \frac{exactitude * recall}{exactitude + recall}$$

L'exactitude et le *recall* sont deux paramètres se calculant sur base du nombre d'occurrences de vrais positifs (VP), de faux positifs (FP) et de faux négatifs (FN) (Tableau 10) selon les formules suivantes :

$$Exactitude = \frac{VP}{VP + FP} ; Recall = \frac{VP}{VP + FN}$$

Tableau 10 : Distribution des vrais/faux positifs/négatifs selon les accords et désaccords des valeurs réelles et prédites par la classification.

		Valeur prédite	
		Accord	Désaccord
Valeur réelle	Accord	Vrai positif	Faux négatif
	Désaccord	Faux positif	Vrai négatif

Coefficient Kappa

L'accord observé entre un ou plusieurs jugements qualitatifs résulte de la somme d'une composante «aléatoire» (hasard) et d'une composante d'accord «véritable» (réel). Cette composante aléatoire gêne l'appréciation que l'on peut avoir de la qualité d'une classification. Afin de contrôler cette dernière, le coefficient Kappa (K) propose de chiffrer l'intensité ou la qualité de l'accord réel. C'est un indice qui permet de «retirer» la portion de hasard ou de subjectivité de l'accord entre les jugements (Bergeri et al., 2002). Il se calcule par la formule suivante :

Tableau 11 : Degré d'accord et valeur de Kappa proposé par Landis and Koch (1977).

Degré d'accord	Valeur de kappa (κ)
Excellent	κ > 0,81
Bon	0,80 < κ < 0,61
Modéré	0,60 < κ < 0,41
Médiocre	0,40 < κ < 0,21
Mauvais	0,2 < κ < 0
Exécrable	κ < 0

$$k = \frac{p_o - p_e}{1 - p_e}$$

où p_o et p_e sont les proportions d'accords observées et aléatoires, respectivement.

$$p_o = \frac{\sum_{i=1}^n m_{i,i}}{M}$$

$$p_e = \sum_{i=1}^n \left(\frac{C_i * G_i}{M} \right)$$

Avec :

- M** Nombre total de points évalués dans la matrice de confusion
- C_i** Nombre de points classifiés dans la classe i (VP + FP)
- G_i** Nombre de points réels de la classe i (VP + FN)

- n** Nombre de classes
- i** Numéro de la classe
- m_{i,i}** Nombre de vrais positifs de la classe i

3.2.3 Logiciels

Le traitement des données fut réalisé à l'aide de plusieurs logiciels :

Logiciel	Version	Description
ArcMap	10.6	ArcMap est la principale application d'ArcGIS. Elle permet d'effectuer un large éventail d'opérations liées aux Systèmes d'Informations Géographiques (SIG). http://desktop.arcgis.com/fr/arcmap/10.3/main/map/what-is-arcmap-.htm
Easy GPS	7.06	Ce logiciel permet la gestion et l'extraction des données de GPS. https://www.easygps.com/
eCognition	10.36.1.8	Utilisé pour de nombreuses fonctions liées à la cartographie de la végétation, détection d'objet, de changements ainsi que l'extraction de caractéristiques par objet. Les multiples processus de segmentations détaillés dans ce document furent réalisés à partir de ce logiciel, considéré comme l'un des plus performant pour l'analyse d'image par objet (OBIA). http://www.ecognition.com
Orfeo ToolBox	6.6.1	L'Orfeo ToolBox (OTB) est un projet open source permettant la réalisation de multiples tâches dans le domaine de la télédétection telles que les opérations de classification, de segmentation ou de traitement de données optiques, multispectrales et radar. Il fut utilisé dans le cadre de cette étude pour réaliser les classifications via l'algorithme Random Forest ainsi que les validations de ces dernières. https://www.orfeo-toolbox.org/
RStudio	1.0.153	Logiciel basé sur le langage de programmation R ayant permis l'analyse statistique et le traitement des données collectées. https://www.rstudio.com/
SNAP	6.0	Sentinel Application Platform (SNAP) est l'interface permettant de réaliser de multiples opérations sur les images provenant des missions <i>Sentinel</i> . SNAP fut utilisé dans le cadre de cette étude pour réaliser les opérations de prétraitement et de traitement des images Sentinel-1. Ces différentes opérations incluent les calibrations, l'intégration des éphémérides, les corrections de terrain, de pente, du bruit d'origine thermique et l'utilisation des multiples filtres speckle. https://step.esa.int/main/toolboxes/snap/

Chapitre 4 : Résultats et discussion

Ce chapitre vise à présenter les différents résultats obtenus via la campagne de terrain ainsi que les opérations de segmentation et classification. Les performances et limitations des classifications seront discutées à la suite de chacun des résultats.

4.1 Campagne de terrain

La campagne de terrain permet de mettre en évidence de nombreuses classes forestières via la collecte de données sur 52 zones. Cependant, des 16 classes initialement identifiées sur le terrain, il fut décidé de regrouper celles-ci en seulement 9 classes, les données radar ne permettant pas de clairement les discriminer. L'entièreté des données acquises sur le terrain est reproduite en annexe.

Forêts mixtes

Il est difficile sur base des images S1 d'identifier le stade de régénération d'une parcelle qui fut l'objet d'une exploitation sélective (Cassol et al., 2019). La Congolaise Industrielle des Bois (CIB) entretient des rotations de 30 ans dans ses concessions et ne sélectionne qu'un nombre limité de grumes ayant atteint un seuil de dimension minimal sur l'ensemble de l'assiette de coupe. Ces coupes sélectives permettent de limiter l'impact des exploitations sur les sols et le régime hydrique tout en facilitant les reprises de végétation. On distinguera donc deux types de forêts, basés sur la récence de leur exploitation : les forêts mixtes anciennement exploitées et les forêts mixtes récemment exploitées.

Forêt mixte anciennement exploitée

La canopée de ces milieux est fermée de manière homogène empêchant la croissance de marantacées et autres essences héliophiles. Ces milieux sont occupés par une grande diversité d'essences.

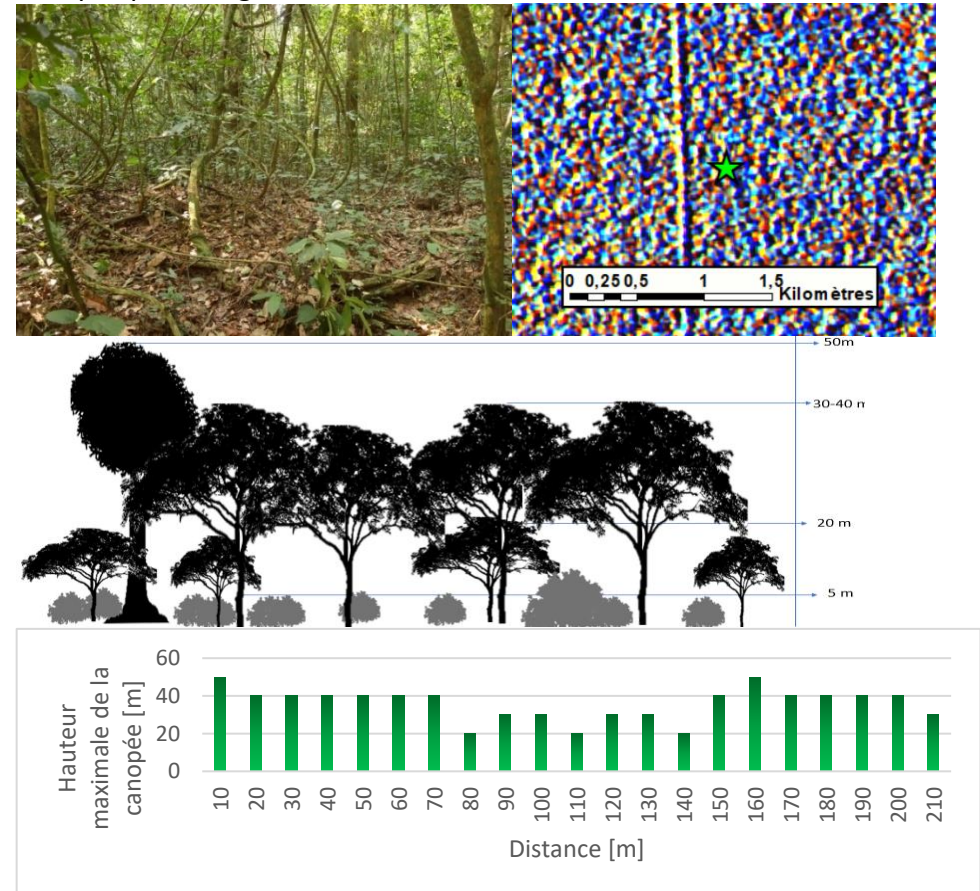


Figure 34 : Visualisation des forêts mixtes anciennement exploitées sur base des données SAR-C [composition VV, VH, VV/VH] (en haut à droite), de terrains (en haut à gauche) et de transect (en bas).

Forêt mixte récemment exploitée

La canopée de ces milieux présente de nombreuses discontinuités. Ces milieux sont ponctués d'anciennes éclaircies généralement recolonisées par des marantacées ou des parasoliers. Ils recèlent également de multiples essences très diverses.

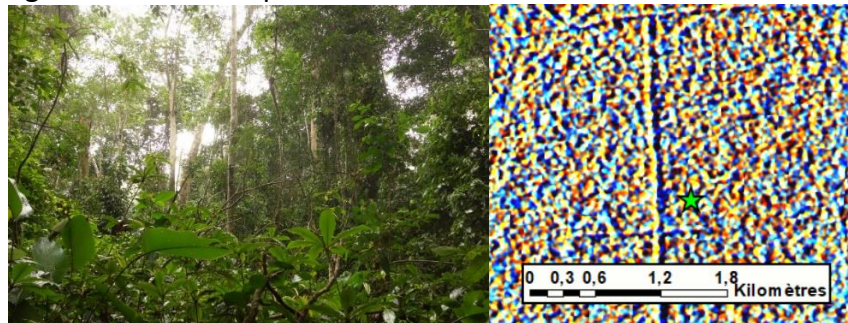


Figure 35 : Visualisation des forêts mixtes récemment exploitées sur base des données SAR-C [composition VV, VH, VV/VH] (en haut à droite), de terrain (en haut à gauche) et de transect (en bas).

Milieux anthropisés – Parasoleraie

Ces milieux, reliquats de parcelles agroforestières, ont été soumis à d'importantes perturbations antérieures et sont rapidement recolonisés par le parasolier (*Musanga cecropioides*), une essence pionnière.

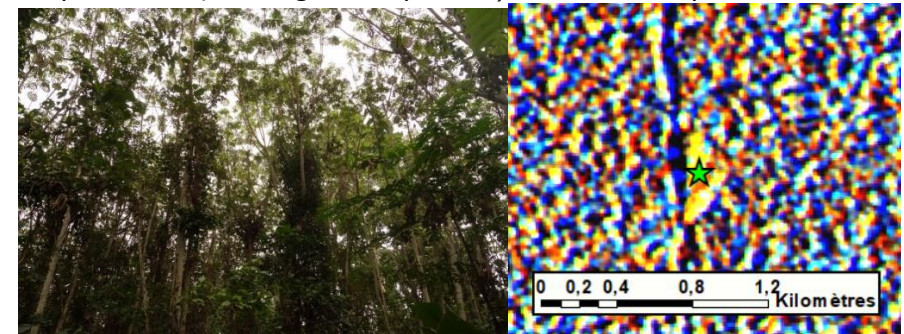


Figure 36 : Visualisation des parasoleraies sur base des données SAR-C [composition VV, VH, VV/VH] (en haut à droite), des données terrain (en haut à gauche) et des transects (en bas).

Forêt marécageuse à Baya aquatiques

Ces marécages, principalement occupés par les bayas et les raphias, présentent une canopée discontinue et clairsemée. Le milieu est inondé tout au long de l'année.

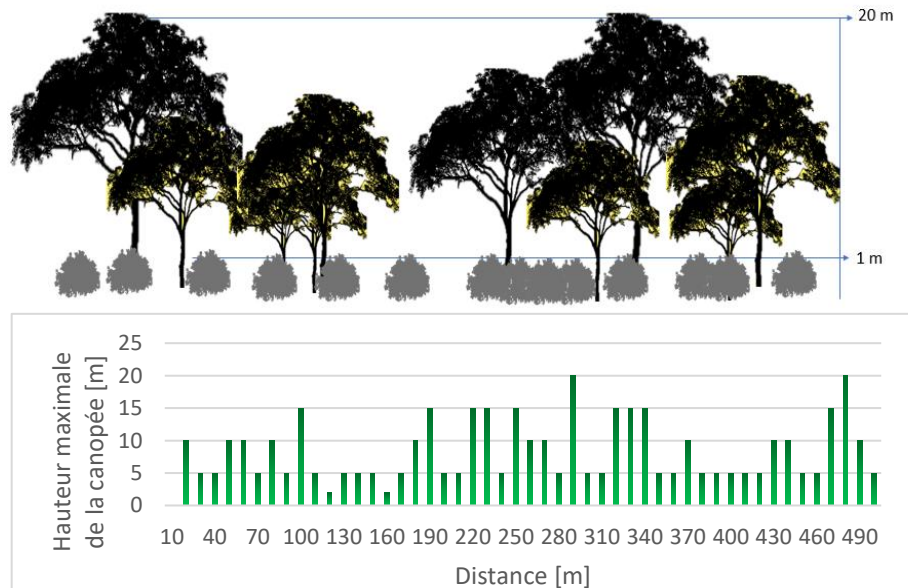
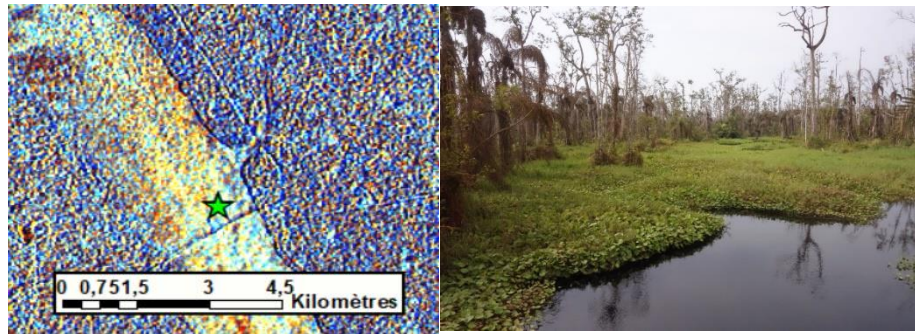


Figure 37 : Visualisation des marécages à Bayas sur base de données SAR-C [composition VV, VH, VV/VH] (en haut à droite, des données terrain (en haut à gauche) et des transects (en bas).

Forêt marécageuse à Raphia – Raphia sp.

Ces forêts sont situées sur des sols hydromorphes présentant de grandes variations du niveau des nappes entre les saisons sèches et humides. Les sols sont tourbeux et imprégnés d'eau en saison sèche alors qu'ils sont entièrement inondés en saison humide. Les raphiales dominent ces forêts.

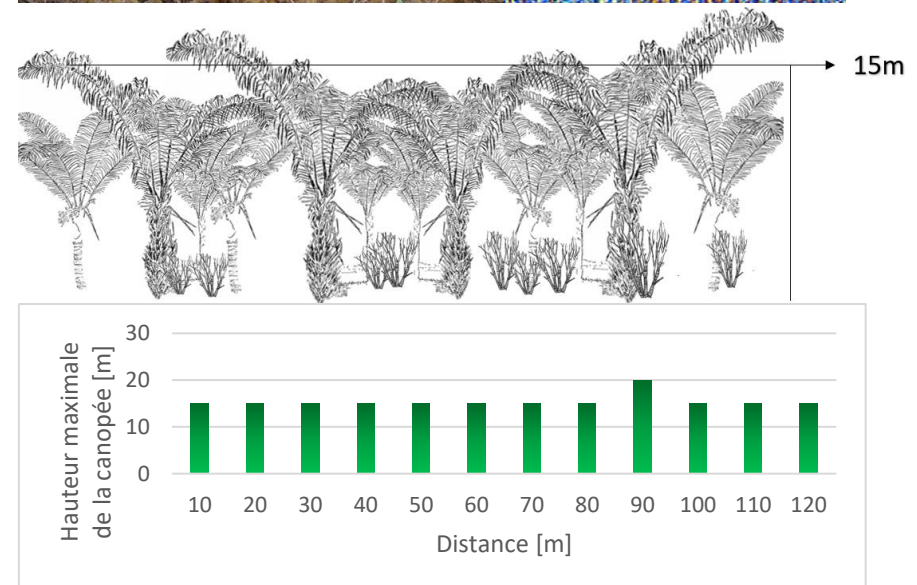
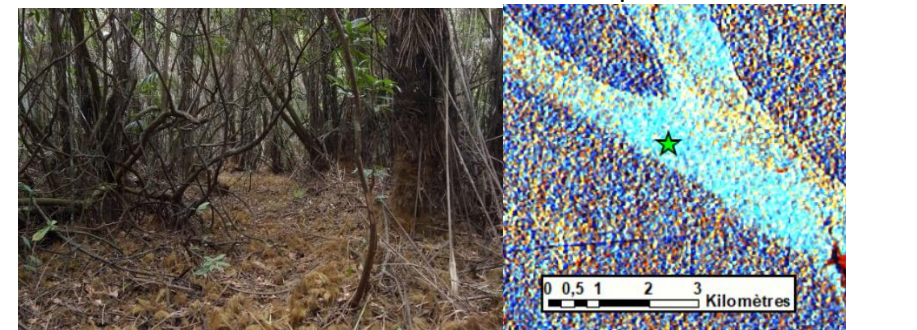


Figure 38 : Visualisation des marécages à Raphia sur base des données SAR-C [composition VV, VH, VV/VH] (en haut à droite), des données terrain (en haut à gauche) et des transects (en bas).

Forêt marécageuse à bubinga – *Guibourtia demeusi*

Ces milieux sont dominés par le bubinga. Peu répandus, ils sont uniquement présents dans certaines zones marécageuses peu profondes à inondation périodique.

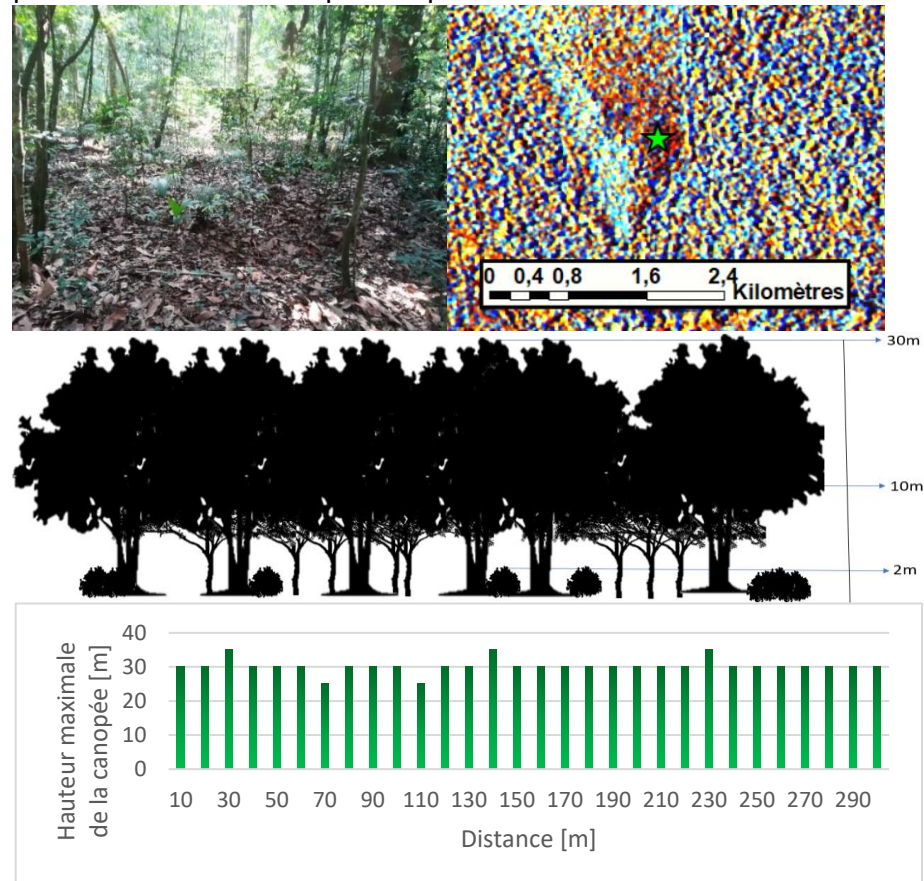


Figure 39 : Visualisation des peuplements de bubingas sur base des données SAR-C [composition VV, VH, VV/VH] (en haut à droite), de terrain (en haut à gauche) et de transect (en bas).

Eyangas

Ces milieux sont caractérisés par des sols saturés en eaux provenant uniquement d'eau de pluie. Certaines de ces zones peuvent être asséchées en saison sèche tandis que d'autres peuvent être sous les eaux tout au long de l'année. Selon le type d'eyanga, la strate arborescente est peu représentée ou absente. Difficilement identifiable à partir des données SAR-C, leur identification fut réalisée sur base des données SAR-L.

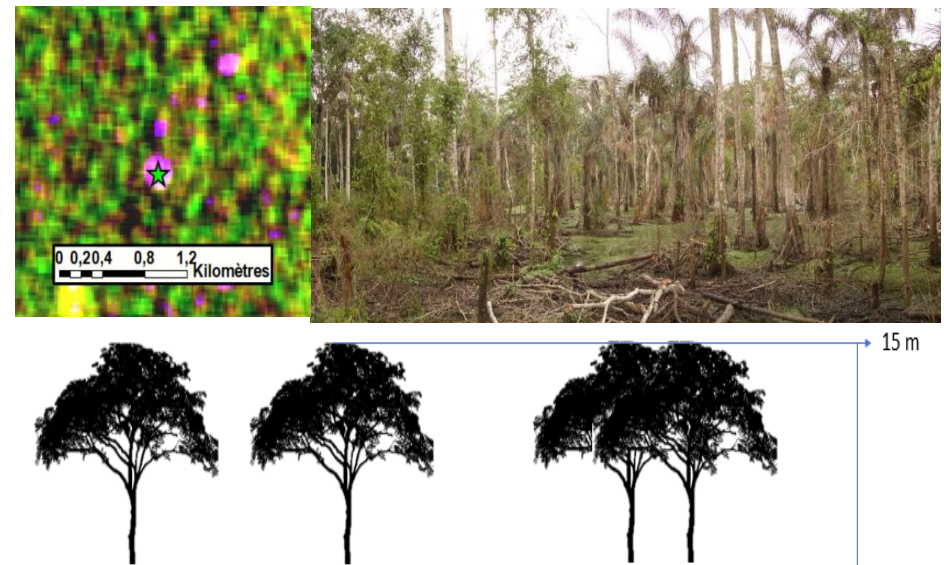


Figure 40 : Visualisation des eyangas sur base des données PALSAR [composition HH, HV et HH/HV], de terrains (en haut) et de transect (à droite). Echelle 1:25.000

Forêt à Limballi – *Gilbertiodendron dewevrei*

Le limballi est une essence pouvant s'adapter à plusieurs conditions hydriques. Il fut retrouvé sur trois types de milieux différents et de manière quasi pure dans la majorité des zones étudiées.

Plateaux (a) :

Ces milieux sont caractérisés par un sol sec et compact et peuvent connaître une faible influence de la saison des pluies se manifestant sous la forme d'inondations de faible amplitude et irrégulièrement réparties sur la surface. Selon les dires des locaux, cet effet saisonnier est de moins en moins marqué d'année en année.

Rives de marécage (b) :

Ces milieux sont secs durant les saisons sèches et se chargent d'eau à la saison des pluies.

Dans les marécages (c) :

Ces milieux sont boueux ou inondés tout au long de l'année.

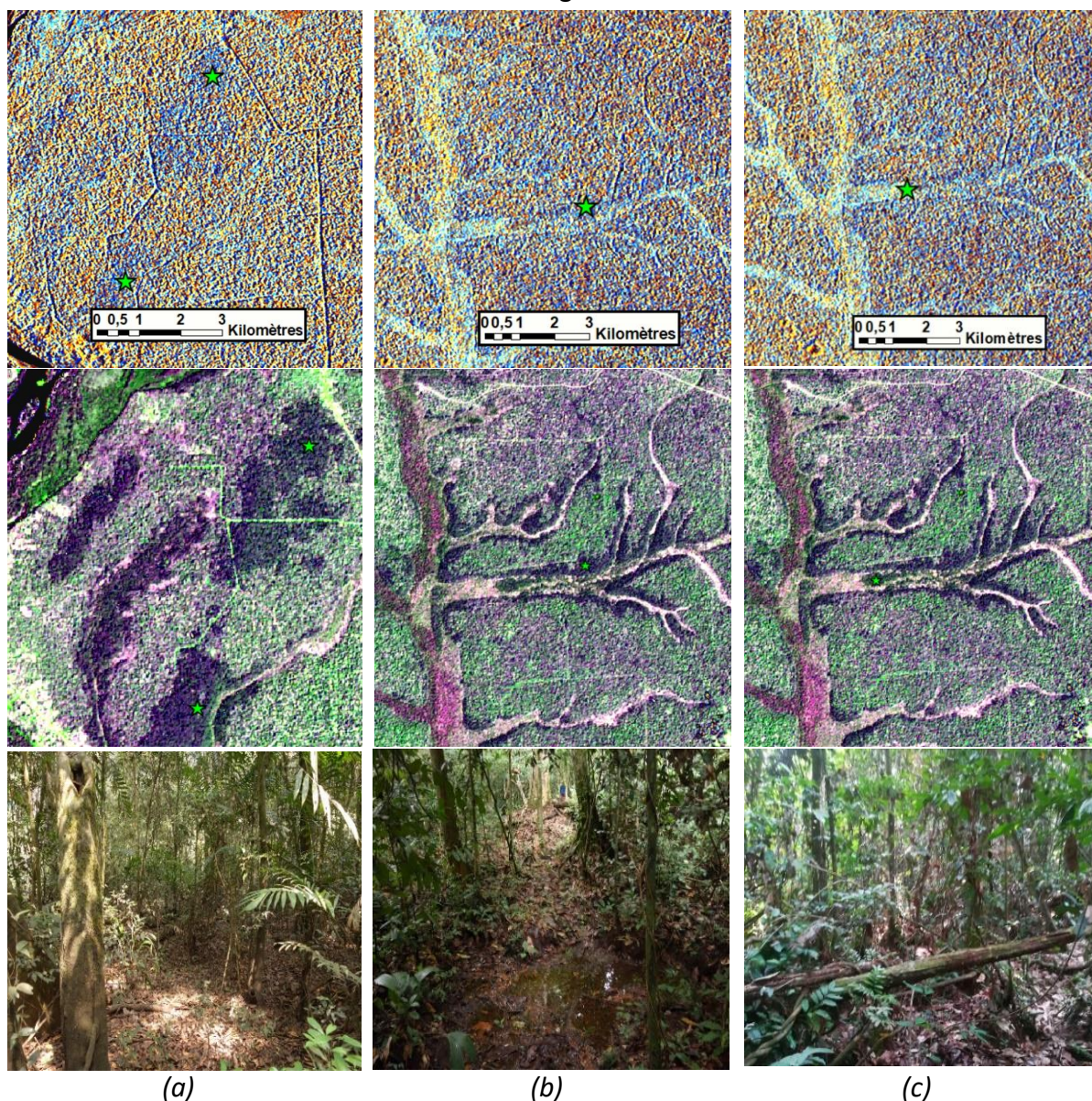


Figure 41 : Identification visuelle des peuplements de limballis à partir des données Sentinel-1 [Composition VV, VH, VV/VH] (au-dessus), Sentinel-2 [bande 6 (rouge), 11 (vert), 8 (bleu) (au milieu)] et sur le terrain (en dessous) selon trois milieux différents : Plateaux (a), milieux rivulaires (b) et milieux marécageux (c).

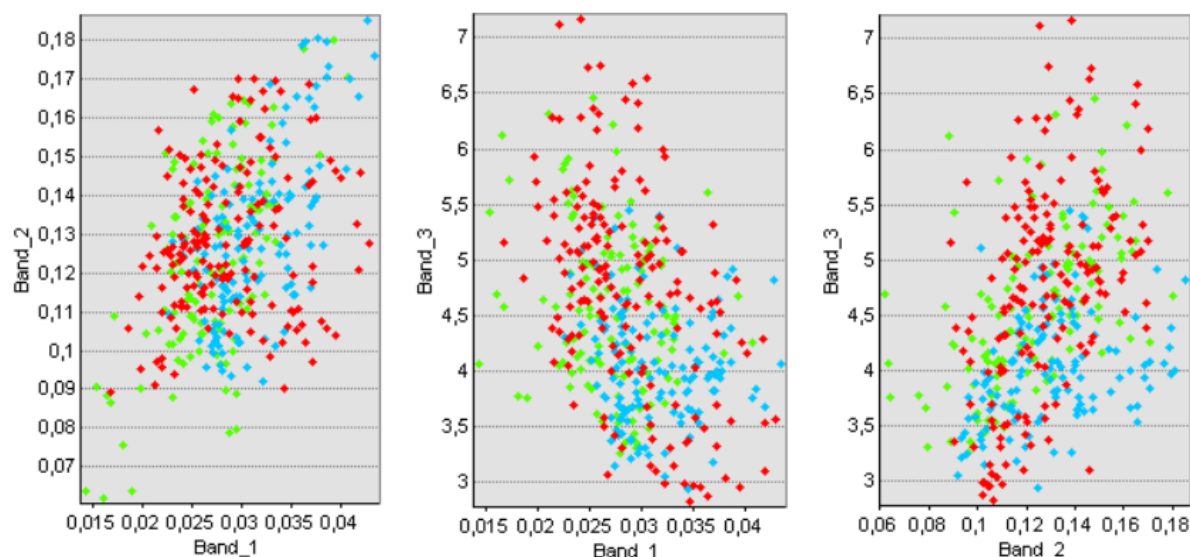


Figure 42 : Distribution des valeurs des coefficients de rétrodiffusion (bande 1 : VV, bande 2 : VH, bande 3 : VV/VH) pour trois classes de peuplements de limballis : plateaux : vert, rivulaires : rouge et marécageux : bleu.

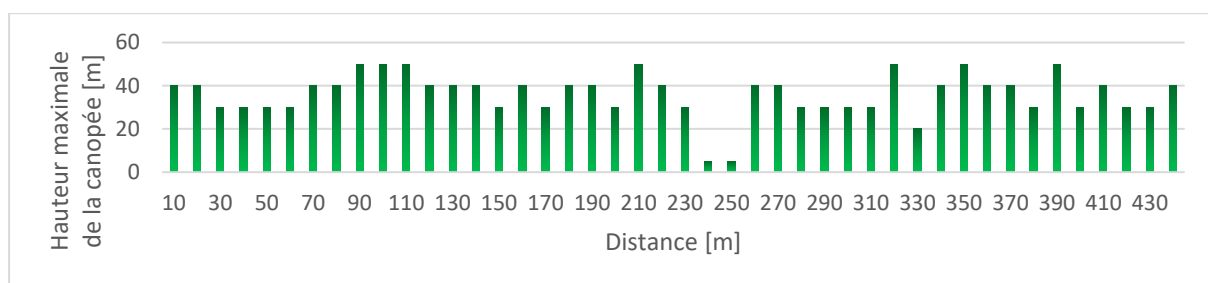


Figure 43 : Données acquises sur la hauteur maximale de la canopée via un transect créé en forêt à limballis

Comme le témoigne la Figure 42, il est difficile de séparer ces trois classes de limballis par simple analyse de leur coefficient de rétrodiffusion.

Prairies inondables

Ces milieux, peu représentés dans la région, ne se trouvent qu'à proximité des marécages. Ils se caractérisent par la présence de nombreuses herbacées se faisant recouvrir d'une nappe d'eau lors de la saison des pluies.

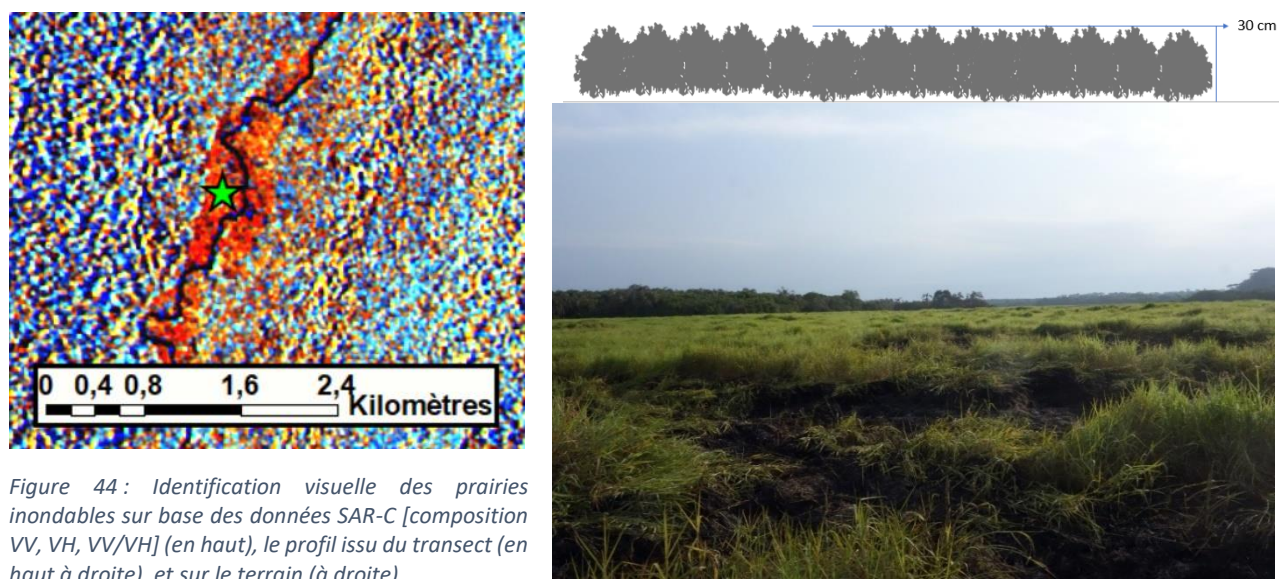


Figure 44 : Identification visuelle des prairies inondables sur base des données SAR-C [composition VV, VH, VV/VH] (en haut), le profil issu du transect (en haut à droite), et sur le terrain (à droite)

4.2 Classification

4.2.1 Classification par pixel

Le Tableau 12 renseigne sur les différentes données ayant servi d'input pour la réalisation des classifications ainsi que les performances respectives de chacune des classifications. Afin de parvenir à une telle classification, 4847 pixels distribués sur 92 zones servirent à l'entraînement de la classification tandis que 638 pixels permirent de réaliser la validation.

Tableau 12 : Ensemble des données utilisées en intrants pour réaliser les différentes classifications par pixels. Les performances respectives de chacune des classifications et classes sont également reprises. Le nombre d'aires d'entraînement de chacune des classes est indiqué entre parenthèses à côté de la classe correspondante.

Données utilisées	Classes	SAR-C [VV;VH;VV/VH]			SAR-C [VV;VH;VV/VH] SRTM [30 m -> 10 m]			SAR-C [VV;VH;VV/VH] SRTM [30 m -> 10 m] Textures [10]			SAR-C [VV;VH;VV/VH] SRTM [30 m -> 10 m] Textures [3]		
		Exactitude	Recall	F-score	Exactitude	Recall	F-score	Exactitude	Recall	F-score	Exactitude	Recall	F-score
	Forêt claire (5)	0,3	0,51	0,38	0,54	0,64	0,59	0,34	0,62	0,44	0,51	0,68	0,58
	Forêt dense (22)	0	0	0	0,74	0,26	0,38	0,94	0,31	0,47	0,73	0,27	0,39
	Forêt à limballis (9)	0,37	0,79	0,5	0,4	0,86	0,55	0,55	0,65	0,6	0,53	0,74	0,62
	Parasoleraie et cultures agroforestières (14)	0,5	0,02	0,05	1	0,27	0,43	0,37	0,43	0,4	0,59	0,5	0,54
	Marécages à raphiales (12)	0,81	0,87	0,84	0,81	0,87	0,83	0,83	0,88	0,85	0,73	0,89	0,8
	Marécages à Baya (7)	0,41	0,72	0,525	0,53	0,81	0,64	0,48	0,86	0,62	0,59	0,88	0,71
	Marécages mixte (9)	0	0	0	0,93	0,15	0,26	0,71	0,05	0,09	0,91	0,3	0,45
	Prairie inondable (4)	0,67	0,87	0,75	0,94	1	0,97	0,84	1	0,91	0,76	1	0,86
	Zone urbanisée (3)	0,782	0,85	0,81	0,67	0,86	0,75	0,81	0,81	0,81	0,6	0,86	0,71
	Eau (3)	1	1	1	1	1	1	0,8	1	0,89	1	1	1
Kappa		0,38			0,52			0,49			0,56		
Exactitude générale		0,46			0,59			0,55			0,62		
F-score		0,5			0,64			0,61			0,67		

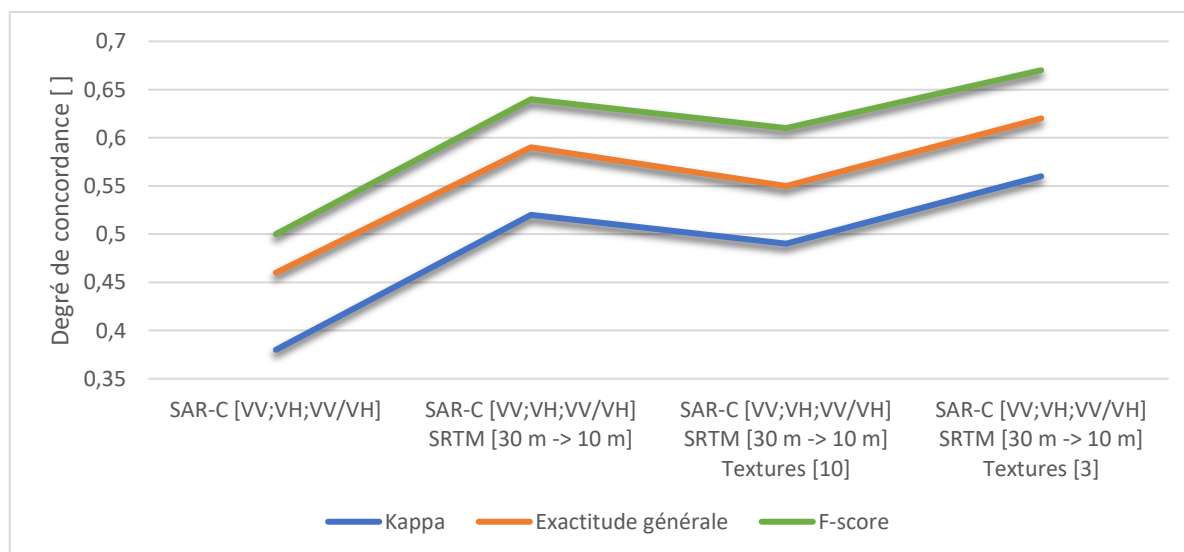


Figure 45 : Evolution du coefficient de Kappa et de la précision générale selon quatre classifications par pixels

On peut observer sur base du Tableau 12 que la simple utilisation des données SAR-C génère de médiocres résultats, deux des classes initialement entraînées n'étant pas représentées sur la classification. L'ajout des données altimétriques SRTM permet l'identification de chacune des classes avec une précision accrue. La discrimination des forêts est cependant mal réalisée, l'algorithme utilisant essentiellement les données topographiques pour les séparer les unes des autres. En conséquence les forêts adultes sont fort représentées sur les zones élevées tandis que les forêts claires sont présentes dans les zones de faible altitude. Cette classification est également polluée par une surabondance de peuplements de limballis sur toute la partie est de l'image.

L'ajout des dix textures calculées sur base des données SAR-C diminue la qualité générale de la classification. Ce phénomène survient lorsque l'entraînement est réalisé sur base d'un trop grand nombre de données. Certaines des données en inputs étant redondantes ou n'apportant pas d'informations pertinentes, elles viennent polluer les résultats des classifications. Néanmoins, les peuplements limballis sont légèrement mieux identifiés sur base de ce jeu de données.

Afin d'éviter la pollution de la classification par la surabondance de données, le set de trois textures (Dissimilarité, Moyenne GLCM et Variance GLCM) dont la corrélation est faible entre chacune des couches fut utilisé comme input pour la classification. L'utilisation de ces données permet d'avoir les meilleures performances de classification.

Cependant, l'utilisation de ces différents paramètres ne permet pas la discrimination des forêts ayant connu de récentes exploitations de celles ayant été exploitées il y a plus de 20 ans. De plus, les exactitudes générales, F-scores et coefficients de kappa de chacune des classifications sont faibles, peu importe le choix des paramètres utilisés pour ces classifications. Ces résultats de faible qualité, sur base la méthodologie appliquée précédemment, témoignent de l'incapacité des données SAR-C à réaliser une discrimination correcte des écosystèmes forestiers via l'analyse polarimétrique de ces dernières. Les matrices de confusions et résultats de chacune des classifications sont reproduits en annexe.

4.2.2 Classification par objet

La classification par objet fut réalisée en utilisant comme input les données SAR-C des coefficients de rétrodiffusion selon les polarisations VV, VH et VV/VH. A ces données furent ajoutées les trois variables texturales liées à la dissimilarité, à la moyenne GLCM et à la variance GLCM ainsi que les données altimétriques SRTM. Le Tableau 13 décrit les différents résultats de la classification par objet. Afin de parvenir à une telle classification, 386 zones servirent à l'entraînement de la classification tandis que 197 permirent de réaliser la validation.

Tableau 13 : Matrice de confusion de la classification mêlant 6 couches objet [VV, VH, VV/VH, Dissimilarité, Moyenne GLCM et Variance GLCM] et le modèle d'élévation de terrain [SRTM].

		Classifié													Recall
		1	2	3	4	5	6	7	8	9	10	11	12	13	
Référence	1	14	1	1	0	0	0	1	0	0	0	0	0	0	0,82
	2	0	14	0	0	0	0	0	0	0	2	0	0	0	0,9
	3	1	4	10	0	0	0	0	0	0	1	0	0	5	0,5
	4	0	1	0	13	0	0	0	1	0	0	0	5	0	0,65
	5	0	0	0	0	14	0	0	0	0	0	0	0	0	1
	6	0	0	0	0	0	12	0	2	0	0	0	0	0	0,86
	7	0	0	0	0	0	0	14	0	0	0	0	0	0	1
	8	0	0	4	0	0	1	0	9	1	0	0	0	2	0,45
	9	0	0	0	0	1	0	0	0	14	0	0	0	0	0,93
	10	0	0	0	0	0	0	0	0	0	13	0	0	0	1
	11	0	0	0	0	0	0	0	0	0	0	15	0	0	1
	12	0	0	0	0	0	0	0	0	0	0	0	10	0	1
	13	0	0	0	0	0	0	0	0	0	0	0	0	8	1
Précision		0,93	0,61	0,67	1	0,93	0,92	0,93	0,75	0,93	0,81	1	0,67	0,53	
F-score		0,87	0,76	0,57	0,79	0,97	0,89	0,97	0,56	0,93	0,9	1	0,8	0,7	

1	Forêt claire
2	Forêt dense
3	Forêt à limballis
4	Parasoleraies et cultures agroforestières
5	Marécages à Bubingas
6	Marécages à Raphias
7	Marécages à Bayas
8	Marécages à végétation mixte
9	Prairies inondables
10	Eyangas
11	Zones urbanisées
12	Sols nus
13	Eau

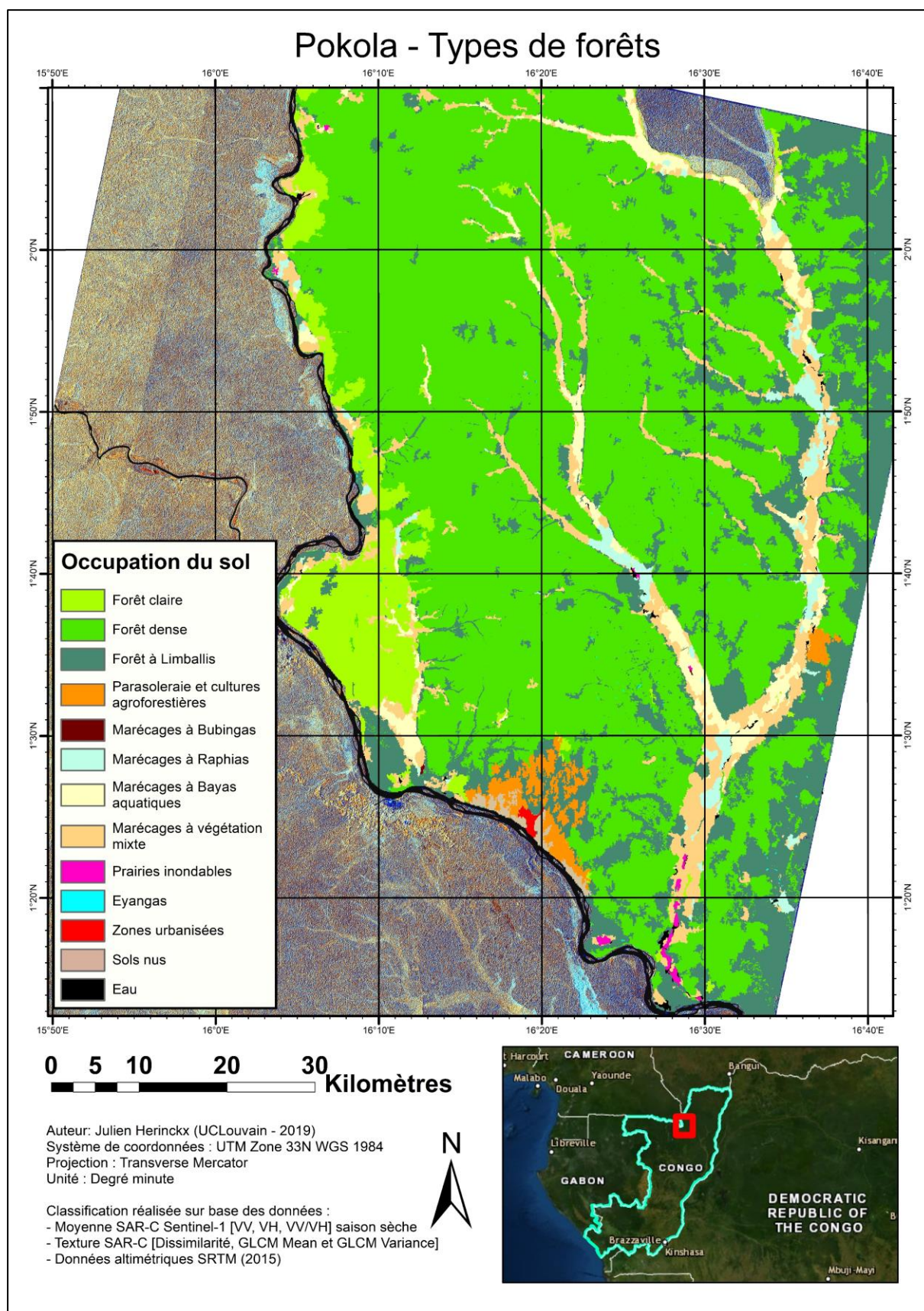
F-score général	0,72
Exactitude générale	0,67
Coefficient de Kappa	0,8

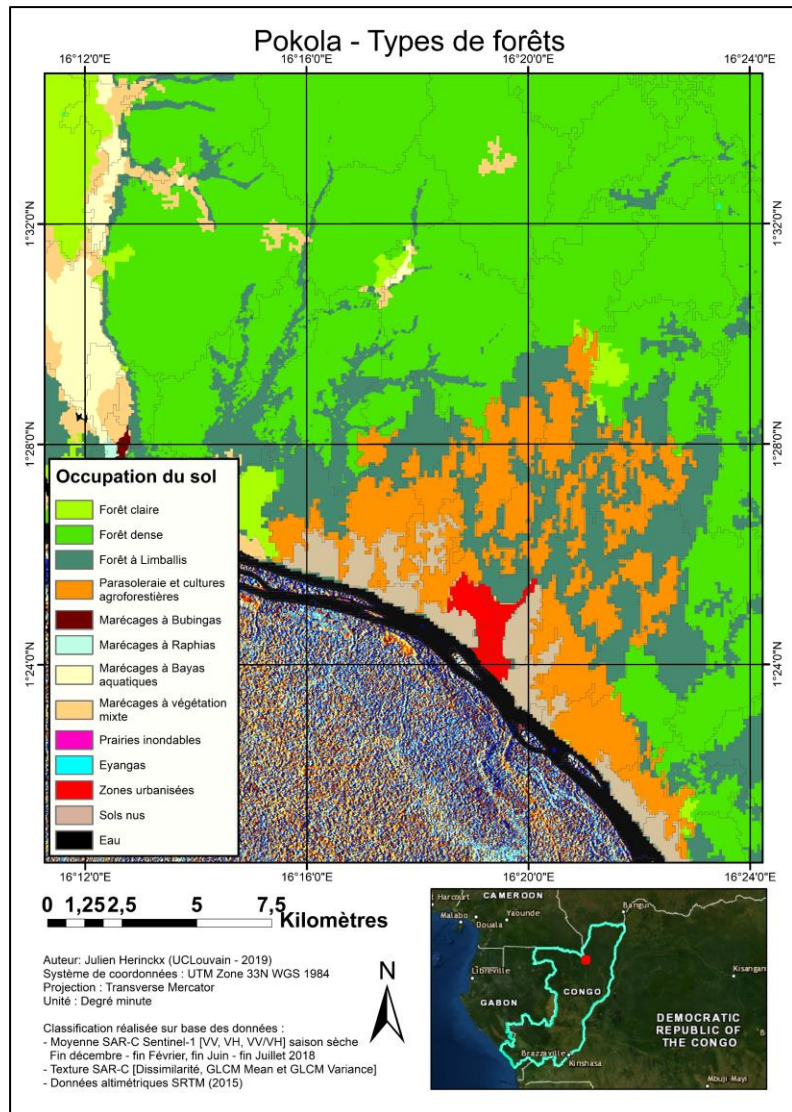
Le Tableau 13 décrit que les zones marécageuses (marécages à bubingas, prairies inondables, marécages à raphiales, marécages à bayas et eyangas) à l'exception des marécages à végétation mixte, présentent des F-scores supérieurs ou égaux à 0,9. Ces valeurs élevées témoignent de la bonne capacité des données SAR-C, en combinaison avec le modèle numérique de terrain (SRTM) à réaliser l'identification des forêts marécageuses. A contrario, les forêts denses et forêts à limballis présentent quant à elles, des résultats de classifications modérés, l'algorithme surestimant les forêts de limballis à l'est de la zone d'étude.

L'ensemble des différentes classes caractérisées par un phénomène hydrologique particulier représente moins de 20 % de la surface totale de la zone étudiée. Les coefficients de kappa, exactitude et F-score généraux, pondérés en fonction de la surface occupée par chacune des classes, sont tirés vers le bas par les classes majoritairement présentes sur le terrain et dont les résultats de classification sont relativement faibles.

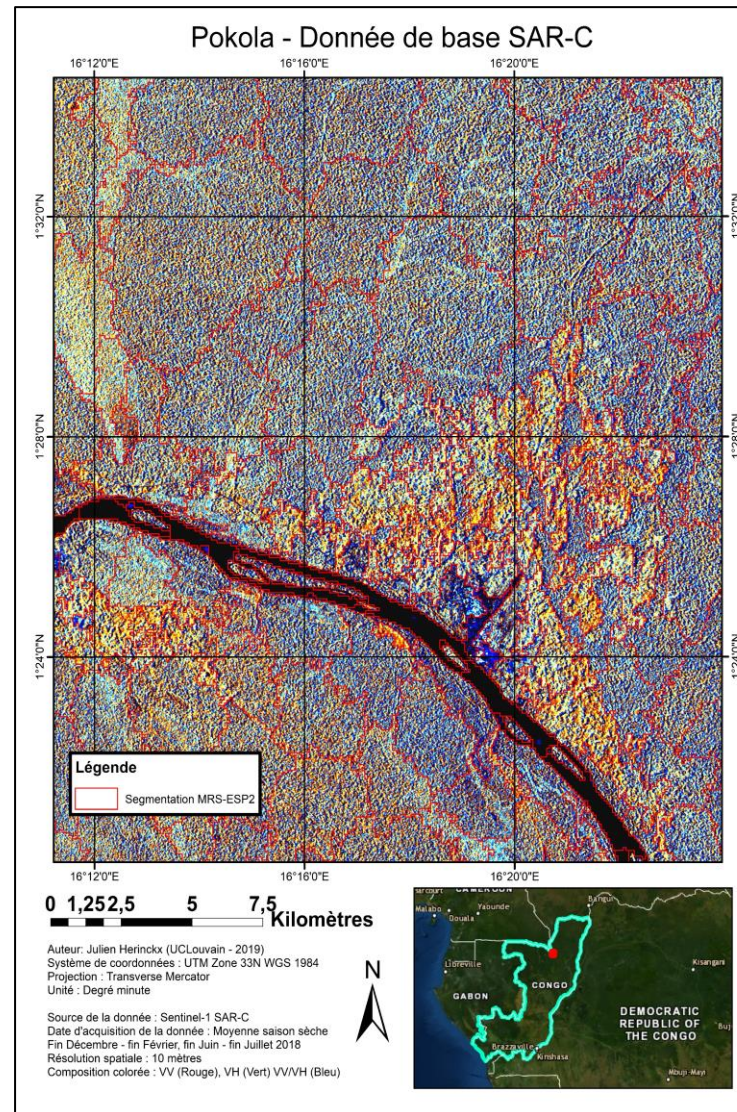
La méthode ayant mené à la classification par objet présentée ci-dessous (Carte 6) est généralisable jusqu'à un certain point. Les opérations de segmentation des zones marécageuses et des forêts sèches sont reproductibles. En revanche, l'extraction de ces dernières peut présenter des difficultés. En effet, les zones marécageuses furent extraites sur base d'un seuil appliqué au paramètre textural de la variance GLCM. Les images acquises étant principalement concentrées sur des milieux forestiers, réaliser une telle extraction à l'échelle d'un pays, où les occupations du sol sont nombreuses et variables, pourrait présenter de nombreuses erreurs.

Afin de mieux interpréter les différents résultats exprimés par le Tableau 13, plusieurs zooms de la classification par objet sont reproduits à la suite de ce document.



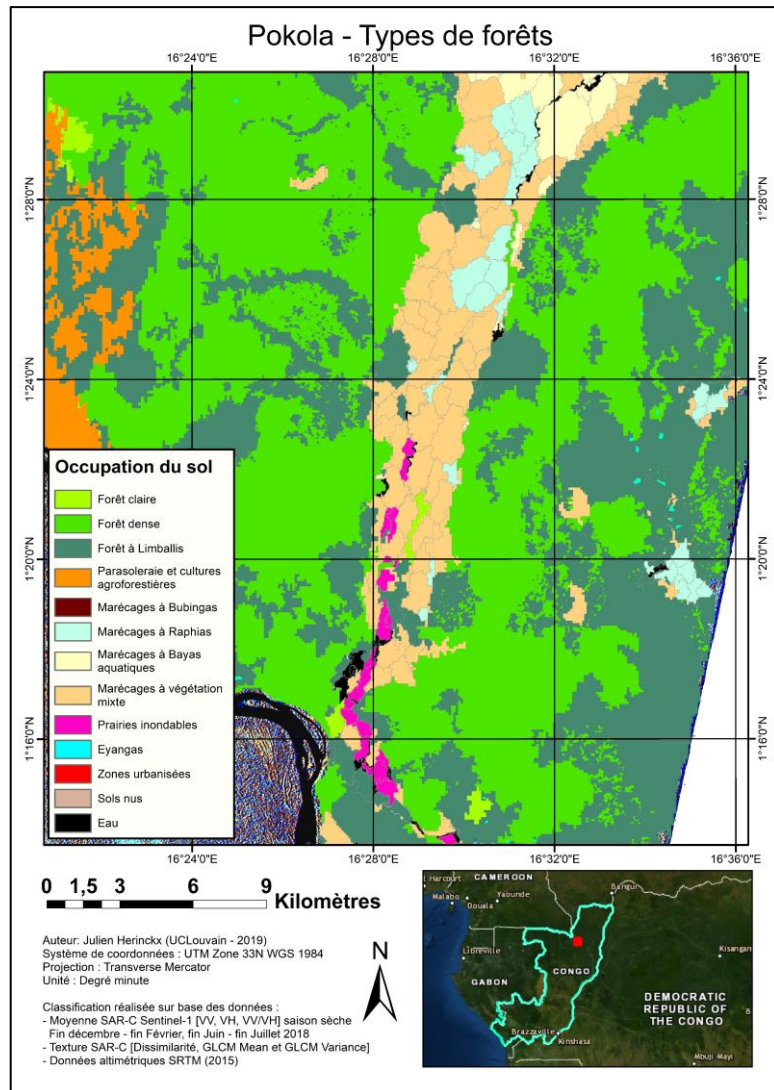


Carte 7 : Zoom de la classification sur les zones anthropisées. La segmentation réalisée sur base de l'algorithme MRS-ESP2 est indiquée par les traits grisés.

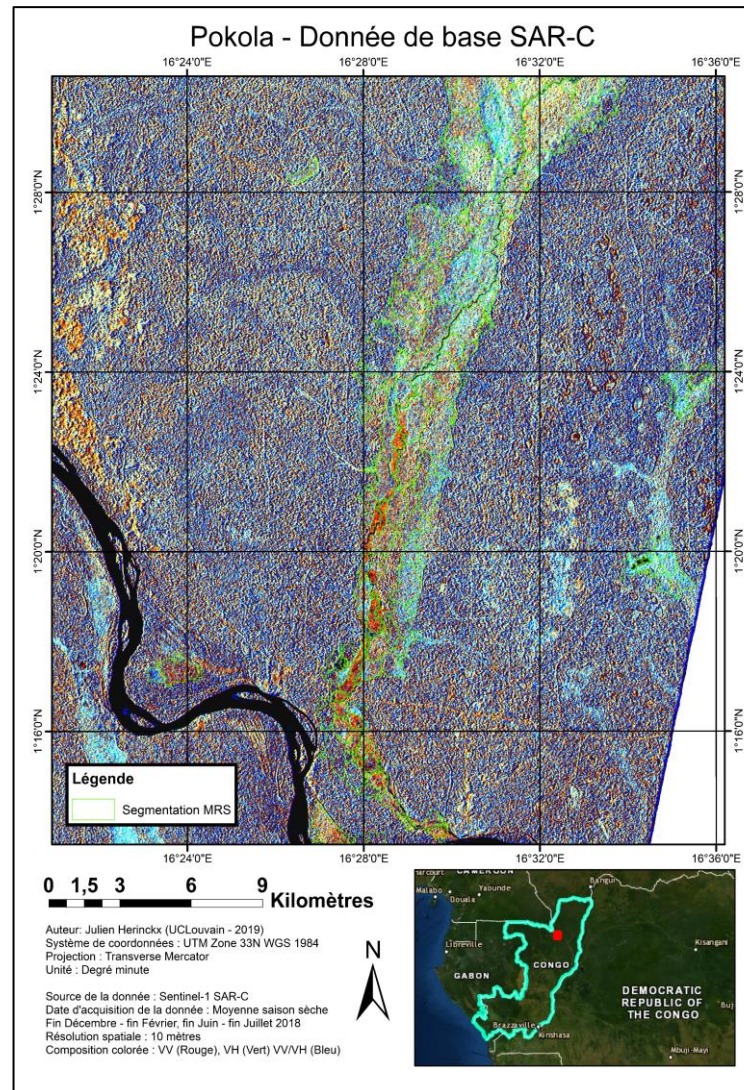


Carte 8 : Zoom sur les zones anthropisées. La segmentation réalisée sur base de l'algorithme MRS-ESP2 est indiquée par les traits grisés.

L'algorithme de segmentation MRS-ESP2 permet la mise en évidence des zones exploitées en tant que cultures agroforestières ou sur lesquelles croissent les parasoliers, et ce, en respectant leurs limites spatiales. Cependant, certaines erreurs d'omission sont à noter, en particulier au nord-est de l'image. Les forêts à limballis situées au nord-ouest des zones anthropisées sont surreprésentées, l'algorithme associant leur présence aux rives des marécages.

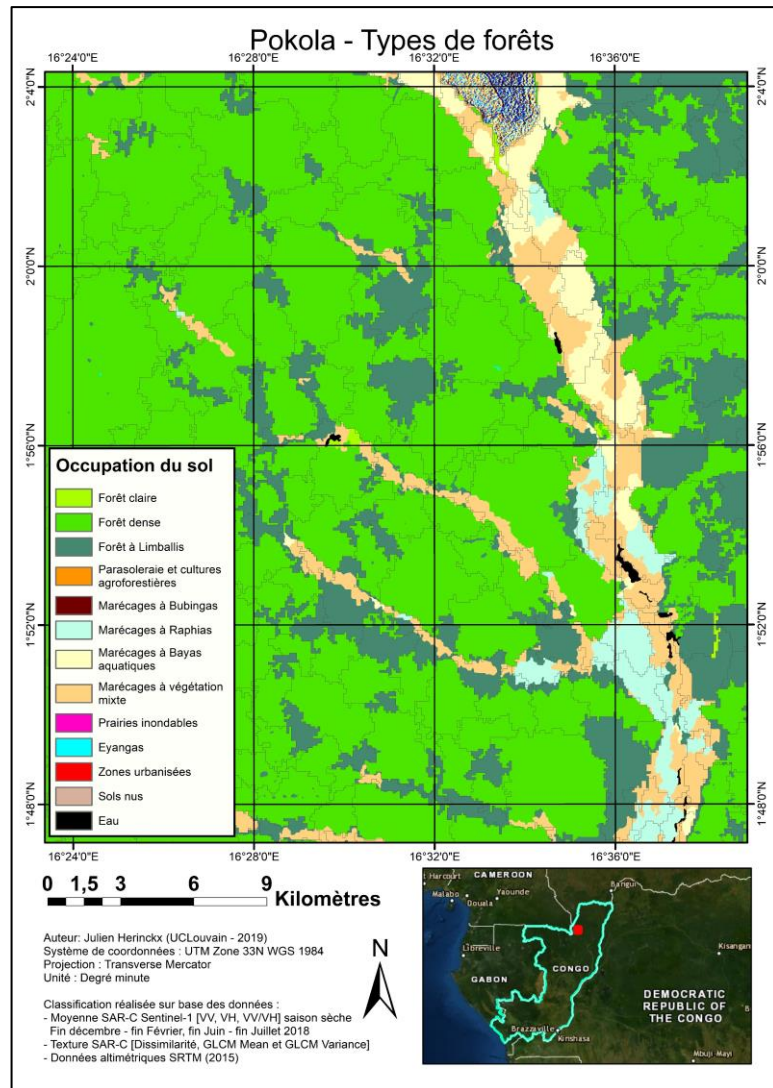


Carte 9 : Zoom de la classification sur les zones marécageuses et les prairies inondables. La segmentation fine des marécages, réalisée sur base de l'algorithme MRS, est indiquée par les traits grisés.

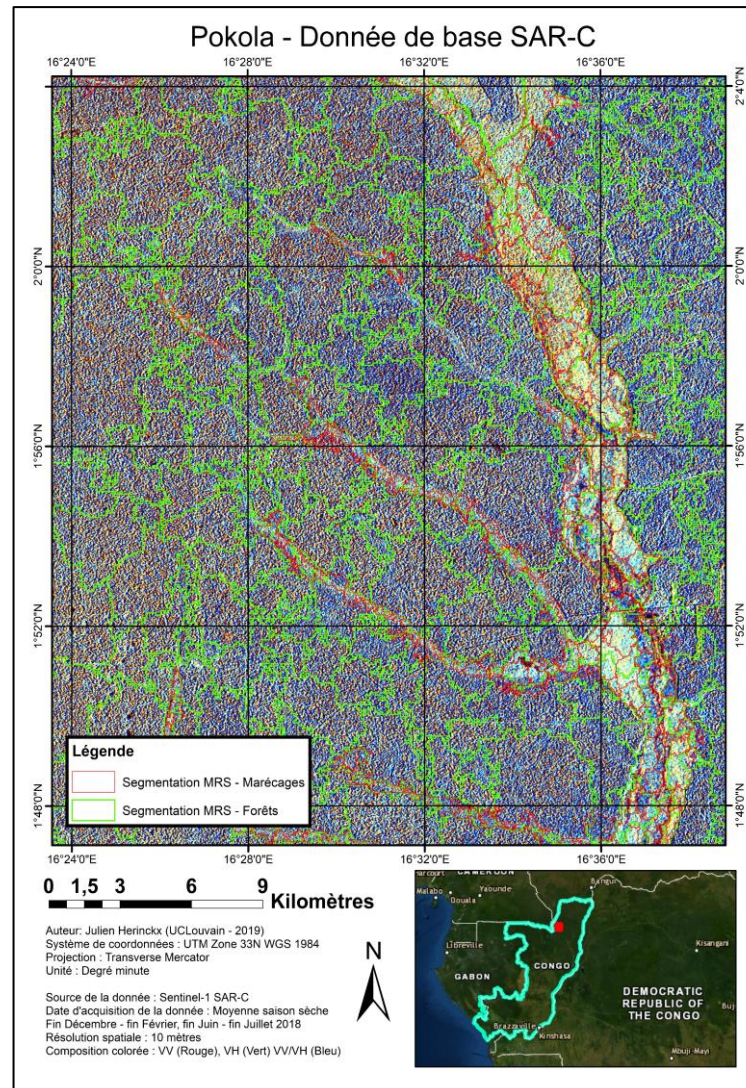


Carte 10 : Zoom sur les zones marécageuses et les prairies inondables. La segmentation fine des marécages, réalisée sur base de l'algorithme MRS est indiquée par les traits verts

Deux segmentations ont été appliquées sur les marécages, la première grossière, servant à extraire ces milieux et la seconde, plus fine, permettant de séparer les différents objets présents au sein des zones marécageuses. La segmentation fine, représentée à l'image permet l'extraction de manière qualitative des différents éléments composant les marécages. Ce n'est pas le cas de la segmentation grossière dont le seuil arbitraire choisi pour réaliser l'extraction omit plusieurs zones telles que celles situées sur la partie est de l'image. Des erreurs de classification sont également à dénoter, et ce, principalement, pour la classe de marécages à végétation mixte dont le F-score est faible (0,56), l'algorithme surestimant leur présence. Cette erreur de classification des zones marécageuses à végétation mixte est notamment visible aux coordonnées suivantes : 16°29'E 1°19'N.

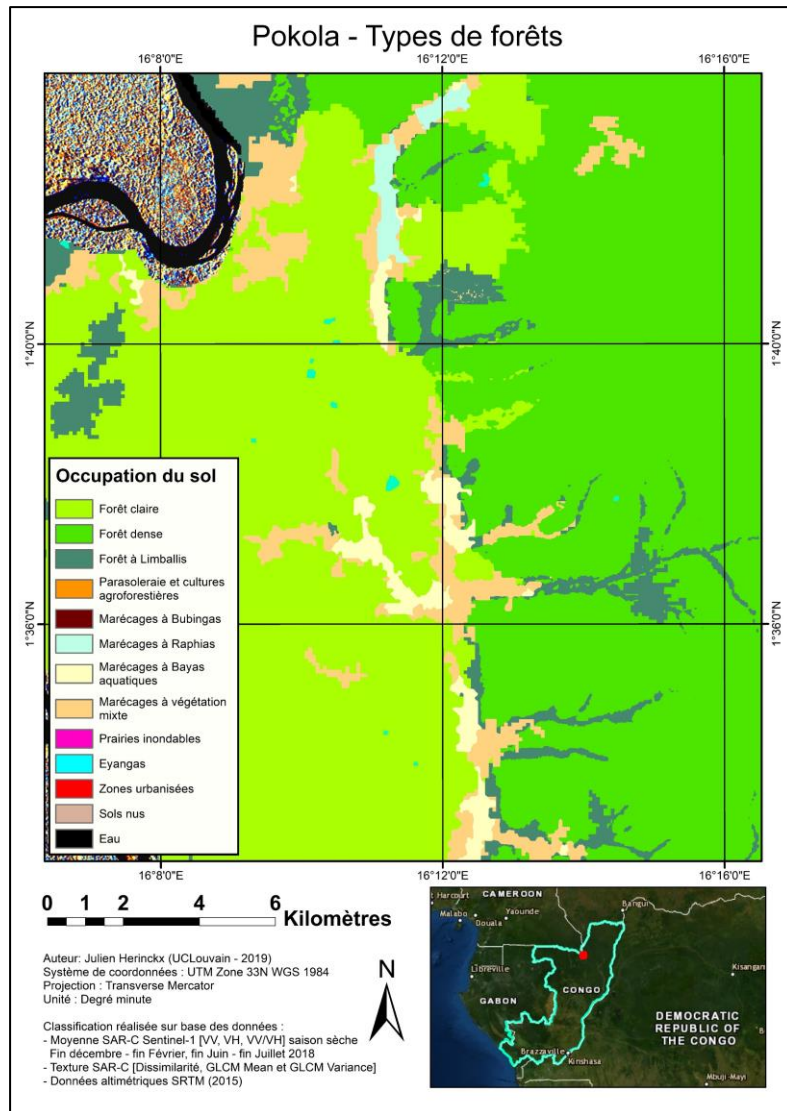


Carte 11 : Zoom sur la classification des forêts sèches et marécageuses. La segmentation, basée sur l'algorithme MRS est indiquée par les traits grisés.

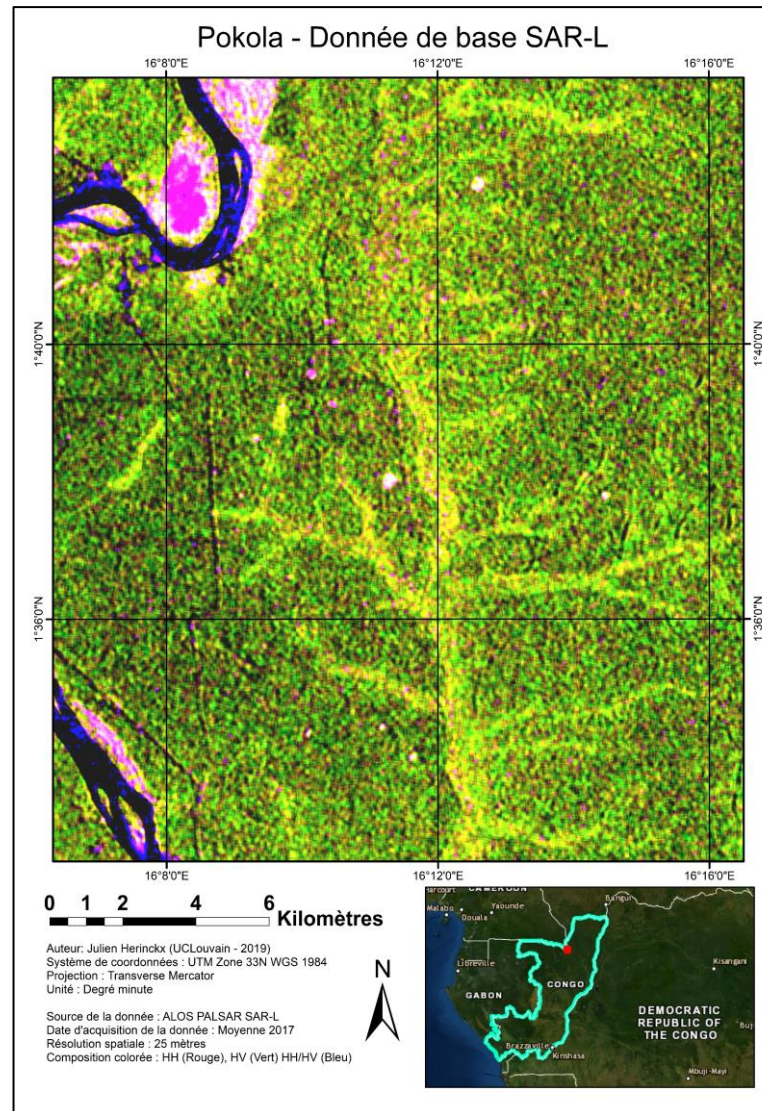


Carte 12 : Zoom sur les forêts sèches et marécageuses. Les deux segmentations ont été réalisées sur base de l'algorithme MRS. Les traits rouges correspondent à la segmentation des marécages et les traits verts à celle de la forêt.

Deux segmentations réalisées sur base de l'algorithme MRS sont représentées à l'image (Carte 12). La première, en rouge, correspond à la segmentation fine réalisée sur les marécages. La seconde, en vert, correspond à la segmentation des milieux forestiers. Les frontières des objets de la segmentation des forêts sont très irrégulières et ne sont pas toujours liées à des limites physiques. On observe une surabondance de peuplements de limballis sur toute la partie se trouvant à l'est de l'image résultant en une classification de qualité modérée avec un F-score de 0,57. La segmentation des marécages, bien que correctement réalisée dans la majorité des cas est soumise à plusieurs erreurs : certains bras de la zone marécageuse n'ont pas été identifiés par l'algorithme tandis que d'autres présentent des discontinuités. De plus, le tracé des cours d'eau au sein des marécages est également mal réalisé.



Carte 13 : Zoom sur la classification des eyangas.



Carte 14 : Zoom sur les eyangas sur base des données SAR-L [Composition colorée : HH, HV, HH/HV]

Les données SAR-L permettent l'identification claire des eyangas qui sont représentés à l'image (Carte 12) selon des tons rosés. Compte tenu de la facilité d'identification de ces milieux, leur classification est correctement réalisée et présente un F-score de 0,90.

Ces images témoignent également de la surestimation des peuplements de limballis à défaut des zones marécageuses dont l'extraction fut incorrectement réalisée.

L'ensemble des résultats observés sur les précédentes cartes permet de mettre en évidence plusieurs limitations.

4.2.3 Limitations

Le Tableau 13 indique que les classes concernant les zones occupées par des étendues d'eau, des forêts de limballis et les zones occupées par une végétation mixte des marécages présentent un F-score inférieur à celui des autres classes. Plusieurs éléments peuvent expliquer cette faiblesse de classification :

Étendue d'eau

Le fleuve Sangha ainsi que les autres cours d'eau bordant l'extrémité ouest des concessions forestières ont été exclus de la classification. La classification était donc focalisée uniquement sur la mise en évidence des cours d'eau présents dans les marécages. La surface des marécages est majoritairement occupée par des essences diverses telles que les marantacées, les bayas aquatiques ou encore les raphiales. Les zones où l'eau est directement visible à la surface sont relativement peu nombreuses, leur signal se confondant avec la végétation avoisinante. Cette confusion du signal rend difficiles les opérations de segmentation ainsi que de classification et induit dès lors de nombreuses erreurs dans l'identification des étendues d'eau.

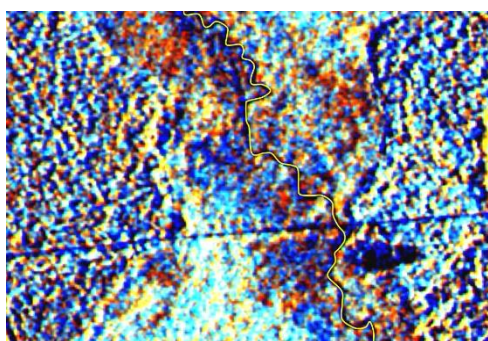


Figure 46 : Visualisation des cours d'eau de la Ndoki sur base des données SAR-C [Composition VV, VH, VV/VH], le tracé du cours d'eau est indiqué en surbrillance (à gauche) et des données terrains (au-dessus). Les images drones ont été capturées le 18 juin 2018 par Laurent Teillier, gérant du groupe Africadrones.

Echelle des données SAR-C 1 :25 000

Marécage à végétation mixte

Des trois types de marécages représentés dans la classification, deux d'entre eux étaient occupés par une végétation homogène caractérisée soit par une dominance de *Raphiales*, soit par une dominance de *baya aquatique* et de *Cleistanthus caudatus*. La troisième catégorie de marécage, quant à elle, n'est pas définie par une végétation dominante particulière. Au contraire, celle-ci est composée d'une végétation mixte d'essences diverses telles que les *Raphiales*, les *Asparagus*, les bayas aquatiques, les *Astonia boneii*, les *Carapa procera* ou encore les *Cleistanthus caudatus*. Outre ce mélange d'essences, les marécages peuvent également être confondus avec des peuplements de limballis qui croissent à même le marécage ou poussent sur les rives de celui-ci. Dans ce dernier cas, les principales erreurs de classification seraient alors dues à des erreurs de segmentation de l'image où l'objet créé ne correspondrait pas exactement aux limites du marécage.

Peuplements de limballis

La Carte 6 témoigne d'une dégradation de la qualité des classifications de l'occupation du sol suite à une surestimation de l'abondance des peuplements de limballis à l'est de la zone d'étude. Cette surabondance de limballi ne représente aucune réalité terrain, mais correspond à un biais radiométrique de l'image. Afin de comprendre l'origine de ce biais, plusieurs pistes furent envisagées. L'ensemble des développements et calculs réalisés dans la suite de ce document sont le fruit d'une recherche personnelle. Les différentes hypothèses et conclusions tirées de ces calculs doivent en conséquence être soumises à vérification avant de pouvoir être validées.

Bien que les interactions des ondes électromagnétiques avec les surfaces d'intérêts soient complexes à décrire, il fut observé par le passé que les phénomènes de rétrodiffusion sont importants lorsque l'angle d'incidence est faible et tendent à diminuer lorsque ce dernier augmente (ASAR Science Advisory Group, 1998). Comme le démontre la Figure 47, la diminution du coefficient de rétrodiffusion est fonction de la rugosité de la surface d'intérêt. On observe une plus faible diminution de ce dernier lorsque la surface présente une rugosité importante. La rugosité d'une surface est dépendante de l'angle d'incidence ainsi que de la fréquence de l'onde émise par le satellite. Cette rugosité peut être qualifiée à partir du critère de Fraunhofer selon la formule suivante :

$$\Delta h < \lambda / 32 \cos \theta.$$

λ = Longueur d'onde [m]
 θ = Angle d'incidence
 Δh = Ecart type des variations de hauteur [m]

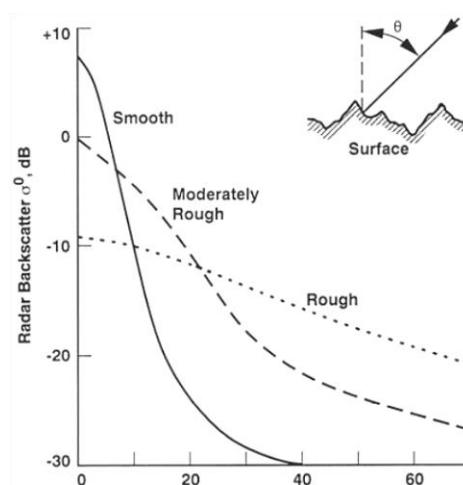


Figure 47 : Variation du coefficient de rétrodiffusion en fonction de l'angle d'incidence d'acquisition des données SAR. La figure démontre un comportement différent de ce coefficient selon la rugosité de la surface d'intérêt. Pour un angle inférieur à 25°, les surfaces modérément rugueuses présentent un coefficient supérieur aux surfaces rugueuses (NASA,s.d). Source de la figure : <https://history.nasa.gov/JPL-93-24/p53b.htm>

Si cette inéquation est respectée, la surface est considérée comme lisse, autrement elle est considérée comme rugueuse. Connaissant la position exacte du satellite (693 kilomètres d'altitude), la largeur totale de la fauchée (250 km), la zone de l'image concernée par le biais radiométrique (20,5 kilomètres) ainsi que les différents angles d'incidence de début (30,07°) et de fin de fauchée (46,06°), il est possible de déterminer l'angle d'incidence exact à partir

duquel commence le biais radiométrique (Figure 49). Une simple relation trigonométrique permet de calculer la distance séparant le satellite du biais radiométrique soit 698,5 km, ce qui correspond presque à l'altitude de l'orbite du satellite. L'utilisation de cette distance pour le calcul de l'arc tangente permet d'identifier l'angle d'incidence correspondant au début de ce biais radiométrique. Ce dernier est égal à $45,2^\circ$. Connaissant les angles d'incidence comprenant la zone de biais radiométrique, on peut, à partir du critère de Fraunhofer, qualifier la rugosité de la surface. L'instrument SAR-C émettant une onde à 5405 MHz, le critère de Fraunhofer indique que la surface est extrêmement rugueuse. On peut dès lors se situer sur le tracé de la Figure 47.

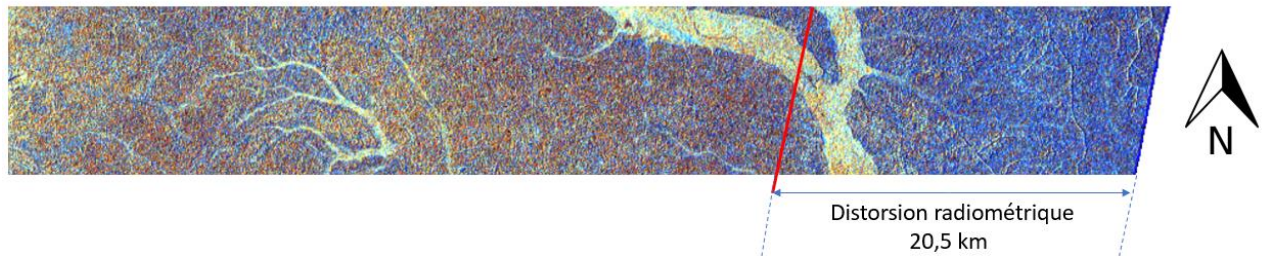


Figure 48 : Visualisation de la distorsion radiométrique présente à l'extrémité de la sous-fauchée IW3 sur une longueur de 20,5 km. Donnée SAR-C [Composition VV, VH, VV/VH].

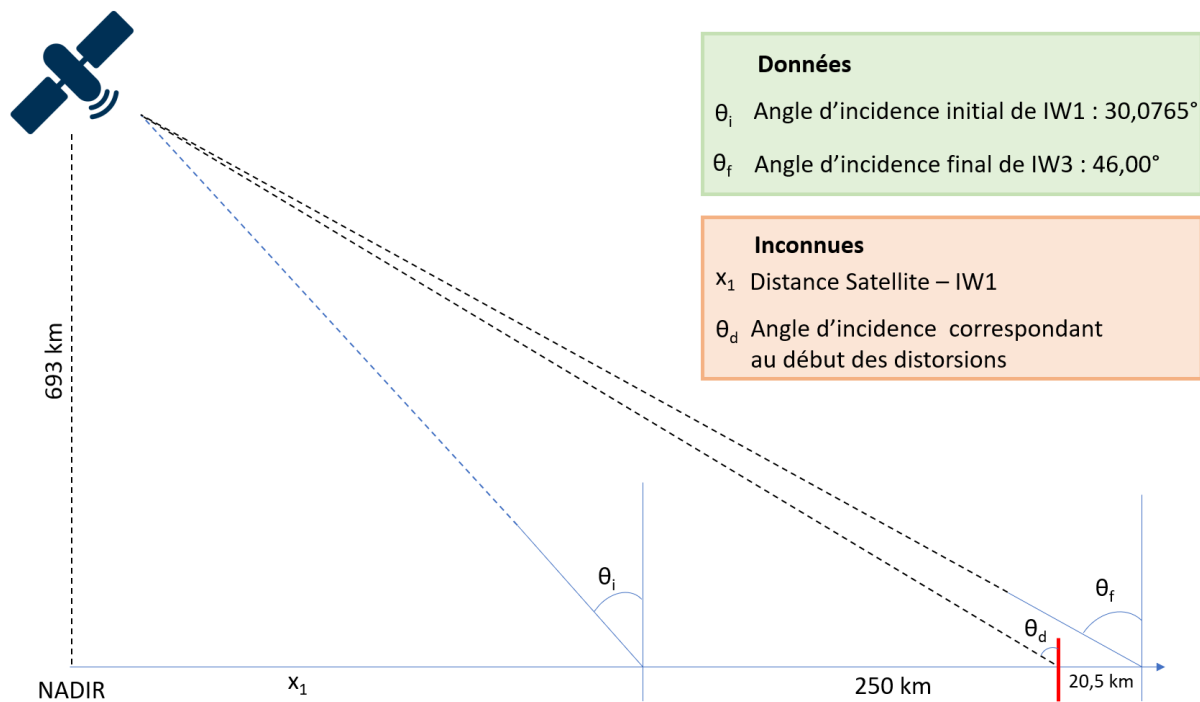


Figure 49 : Schématisation du problème permettant de calculer l'angle d'incidence correspondant aux distorsions radiométriques

De manière à quantifier les variations du signal, un transect de 77 kilomètres, découpé par tronçons de 1 kilomètre fut appliqué à l'image. Sur ces derniers ont été extraites les valeurs moyennes des coefficients de rétrodiffusion. Afin de mieux observer l'effet de gradient des distorsions radiométriques, les données sont approximées par une régression polynomiale de troisième degré. La Figure 50 renseigne sur les différentes valeurs extraites des coefficients de rétrodiffusion pour la polarisation VV, VH et VV/VH.

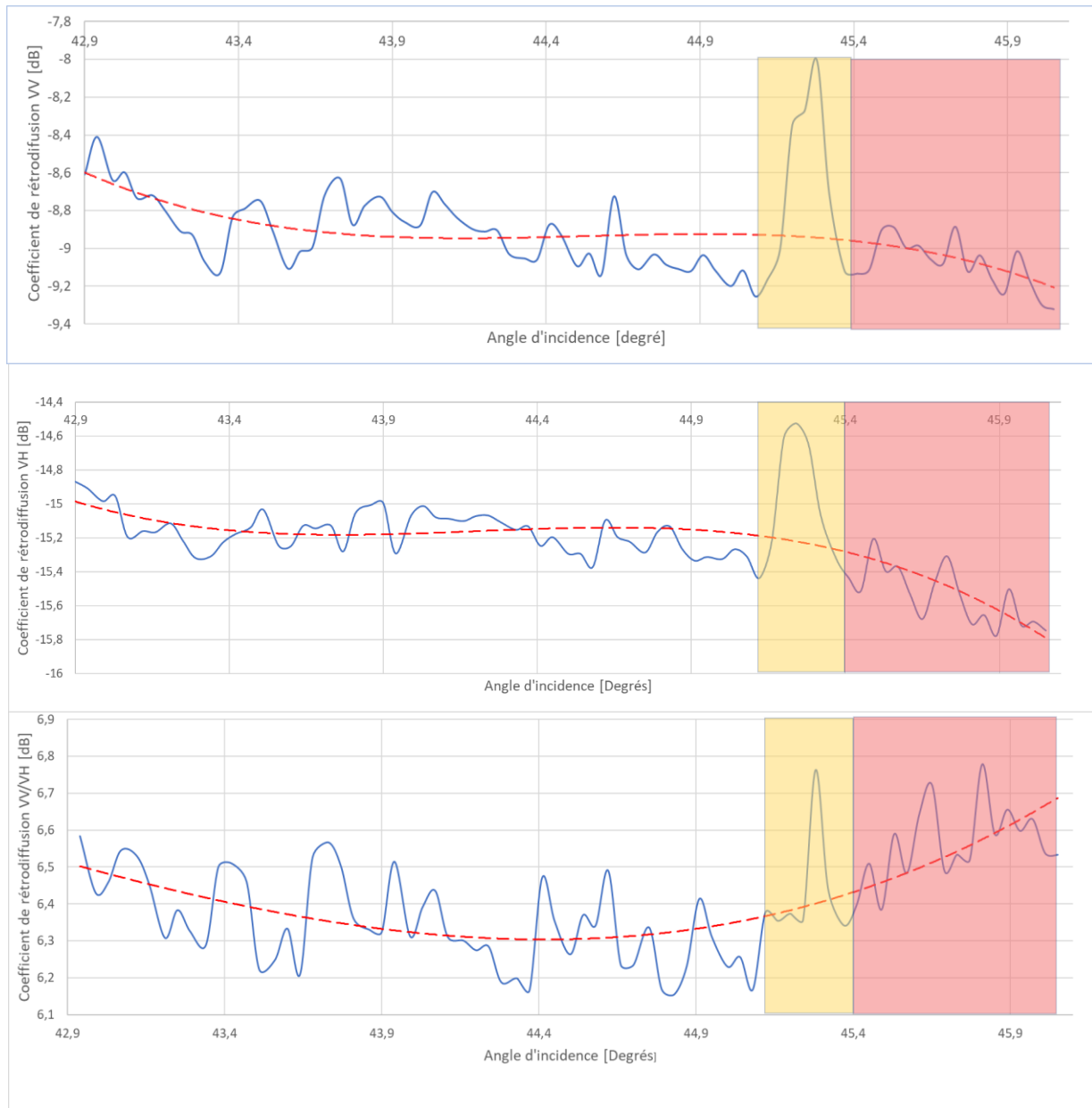


Figure 50 : Variation du coefficient de rétrodiffusion des polarisations VV, VH et VV/VH en fonction de l'angle d'incidence. Ces données sont approximées à partir d'une polynôme de troisième degré. La partie de la courbe mise en surbrillance jaune correspond au signal du zone marécageuse tandis que celle en surbrillance rouge correspond au biais radiométrique.

La Figure 50, conformément à ce qu'indique la Figure 47, décrit que le coefficient de rétrodiffusion diminue de manière quasi linéaire avec l'angle d'incidence. Cependant, cet effet est d'autant plus marqué une fois passé un angle d'incidence de 45,5 degrés. La littérature n'expliquant pas un tel phénomène, il est nécessaire de réaliser de plus amples investigations sur le sujet.

Impact de la pente

Afin d'observer l'impact potentiel de la pente sur les variations radiométriques représentées à la Figure 50, les données d'ombrage et d'exposition des pentes ont été extraites sur base des données SRTM à 30 mètres de résolution.

L'ombrage est une représentation en trois dimensions de la surface d'intérêt prenant en compte la position relative du soleil pour l'application de ladite fonction (ArcMap 10.3, 2016).

Cette fonction fut utilisée de manière à comprendre la manière avec laquelle le satellite perçoit la surface d'intérêt. En conséquence, la position relative du soleil correspond à celle du satellite lors de l'acquisition de la donnée. Ce dernier fut dès lors placé selon un angle d'incidence de 45,5 degrés et un angle azimutal de 282 degrés (0 degré correspondant au Nord, l'angle augmentant dans le sens horloger). L'utilisation de cette fonction ne permet cependant pas de modifier l'altitude de l'hypothétique soleil/du satellite. En conséquence, le soleil est placé à une distance de la surface d'intérêt tendant vers l'infini. Cela permet d'avoir un angle d'incidence constant sur toute la surface d'intérêt, ce qui ne correspond toutefois pas à la réalité observée par le satellite. L'exposition reprend l'orientation des différentes pentes de l'image et les distribue selon quatre points cardinaux et leurs subdivisions.

Afin de faire correspondre ces deux données avec l'image de base, ces dernières ont été rééchantillonnées à 10 mètres de résolution par interpolation bilinéaire pour les données d'ombrage et par la méthode du plus proche voisin pour les données d'exposition.

La combinaison de ces deux données permet de mieux interpréter les variations topographiques de la zone d'intérêt. Celles-ci ressortent comme étant très irrégulières, ponctuées par de nombreux cours d'eau traversant la zone d'ouest en est et causant des alternances de pentes exposées au nord et au sud. Le biais radiométrique se manifestant comme une décroissance progressive du coefficient de rétrodiffusion, il semble difficile de faire corrélérer ce biais avec les effets de la topographie.

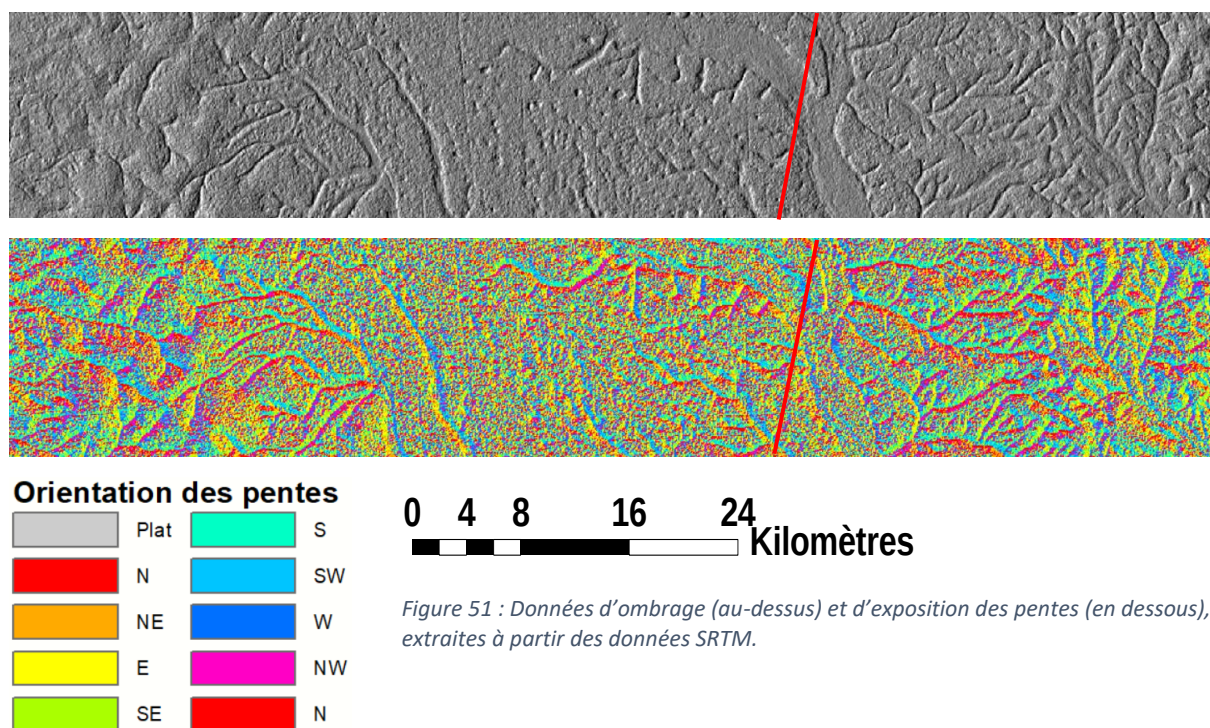


Figure 51 : Données d'ombrage (au-dessus) et d'exposition des pentes (en dessous), extraites à partir des données SRTM.

Impact du prétraitement – Réduction du bruit d'origine thermique

Ce traitement consiste en la réduction du bruit généré par l'augmentation de la température du capteur du satellite suite à l'acquisition de l'image. Ce bruit est plus fréquemment observé dans les milieux présentant de faibles coefficients de rétrodiffusion (inférieur à -20 dB) tels que les étendues d'eau, et ce, principalement pour les polarisations croisées (VH dans le cas de Sentinel-1). Lorsque le satellite acquiert des données sous forme de multiples sous-fauchées, le bruit thermique présente des intensités différentes pour chacune des sous-fauchées. Cette différence d'intensité génère des distorsions radiométriques lorsque les

extrémités des sous-fauchées se chevauchent (Piantanida, 2017). La distorsion présentée à la Figure 48 n'est pas située à la frontière de sous-fauchées se chevauchant, ce phénomène n'explique donc pas la diminution du signal concernant cette zone.

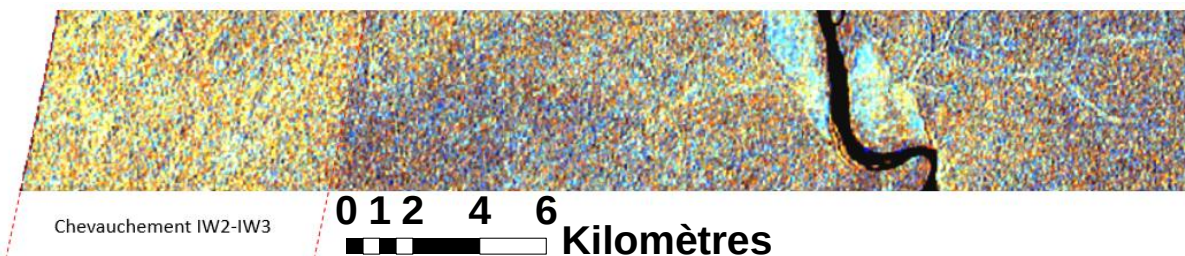


Figure 52 : Visualisation de la variation radiométrique due à la superposition des fauchées IW2 et IW3

Toutefois, le bruit thermique influence également la radiométrie d'une image même lorsque ce signal est supérieur à -20 dB. Influencé par l'angle d'incidence, ce dernier peut introduire un biais allant jusqu'à 3 dB en extrémité de IW1. La Figure 53 démontre l'évolution de ce biais pour chacune des sous-fauchées, mesuré en fonction de l'équivalent bruit de sigma zéro (NESZ) sur la surface équivalente radar. La surface équivalente radar correspond au ratio de l'énergie rétrodiffusée sur l'énergie que le capteur aurait effectivement acquise si la surface cible avait diffusé cette énergie incidente de manière isotropique, le tout généralement exprimé en décibel (Mahafza, 2000). NESZ, quant à lui, correspond au bruit thermique exprimé sur base de la surface équivalente radar (Piantanida, 2017).

La Figure 53 témoigne de l'importante augmentation du biais radiométrique lié au bruit thermique à l'extrémité de chacune des sous-fauchées. Les données des coefficients de rétrodiffusion en polarisation VH indiquent une décroissance du signal de 0,4 dB, ce qui correspond au biais radiométrique des valeurs de l'IW3 proches de -15 dB. Le biais radiométrique présent sur les images serait alors dû à des erreurs de correction du bruit thermique. Ces erreurs de correction pourraient provenir de la mise à jour du 13 mars 2018 de la S1 Instrument Performance Facility (IPF) lorsque cette dernière fut updatée à la version 2.90. Le passage à cette nouvelle version engendra de nombreux problèmes de correction du bruit thermique, et ce, suite à une modification de la manière dont sont encodés les vecteurs de bruits dans les métadonnées, ces derniers étant nécessaires pour réaliser la correction (ESA, 2018; ESA STEP Forum, 2018). Afin de confirmer ou d'infirmer ces hypothèses, de plus amples investigations, dépassant le cadre de ce mémoire, doivent être menées.

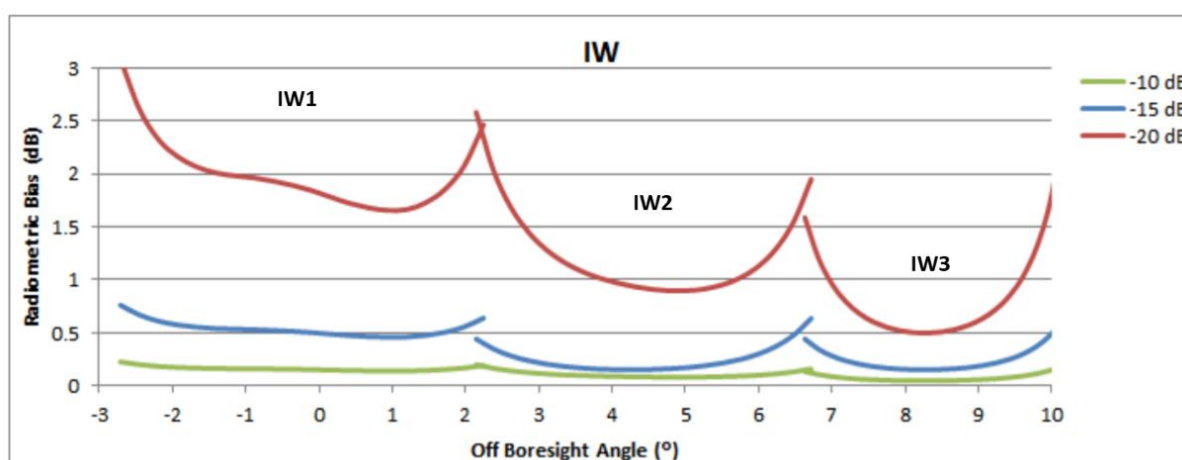


Figure 53 : Biais radiométrique sur la surface équivalente radar (RCS) mesurée en fonction de NESZ (Piantanida, 2017)

4.3 Proposition d'une nouvelle méthodologie

En considération des différentes analyses préliminaires et résultats obtenus aux sections précédentes, la méthodologie suivante est proposée. Celle-ci permet l'obtention de la classification par objet présentant les meilleures performances.

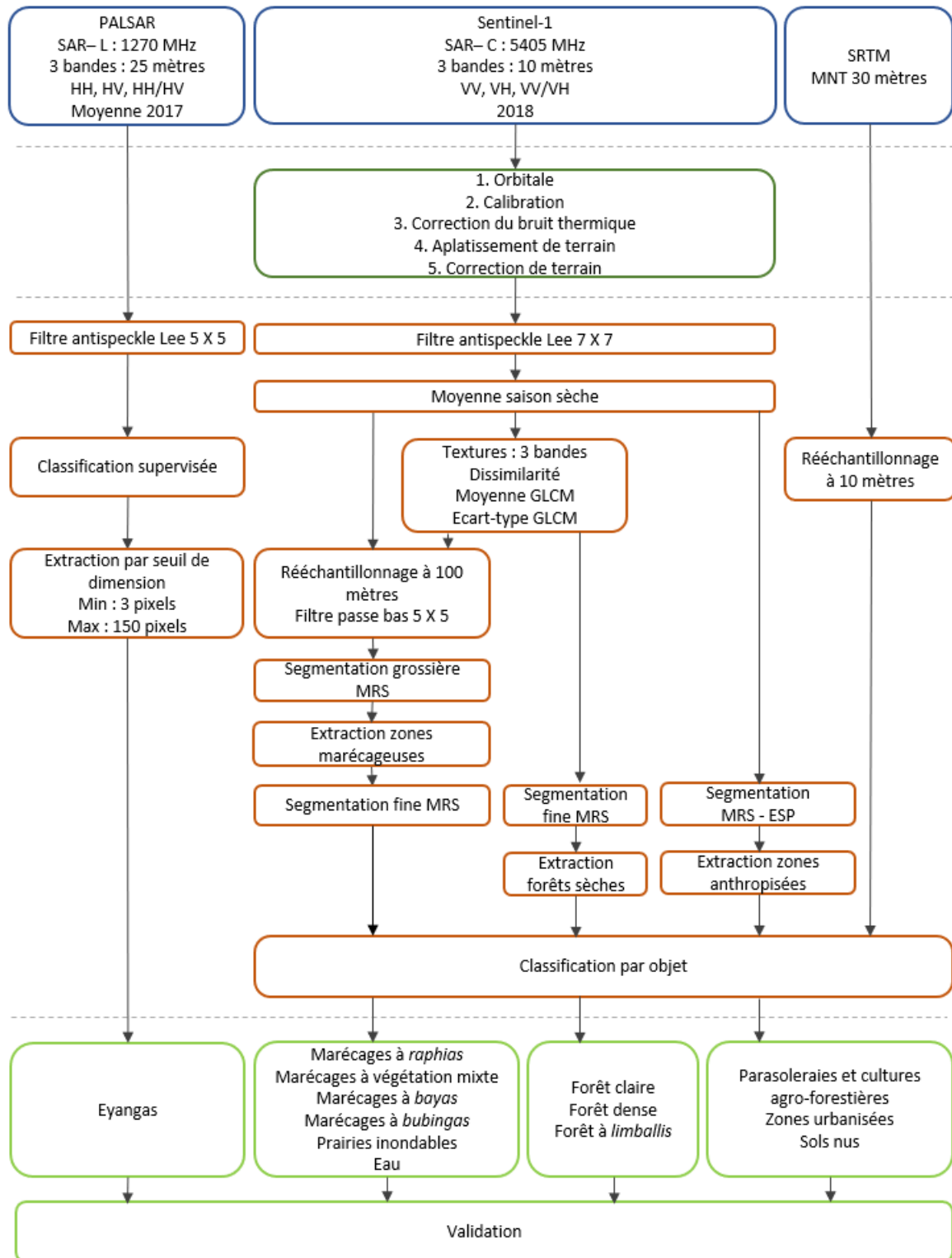


Figure 54 : Proposition de traitement des données permettant l'élaboration de la classification par objet présentant les meilleures performances

Chapitre 5 : Conclusion

Demandé par l'Observatoire des Forêts d'Afrique Centrale et la Commission des Forêts d'Afrique Centrale, ce mémoire s'inscrit dans une dynamique de recherche visant à soumettre des propositions de nouvelles méthodes / combinaisons de méthodes permettant une classification affinée des écosystèmes forestiers des forêts tropicales du bassin du Congo.

Les différents processus de recherche ayant mené à la classification finale ont pu mettre en évidence plusieurs éléments :

Les classifications des données Sentinel-1 par pixel témoignent d'une insuffisance en termes d'exactitude générale à la classification des écosystèmes forestiers, et ce, peu importe les couches mises en input. Cependant, il fut observé que l'ajout du modèle d'élévation DEM permet une augmentation importante de la qualité des classifications des milieux marécageux. A contrario, l'ajout des données texturales Sentinel-1 ont eu pour principal effet la réduction de la qualité générale des classifications. Cette diminution de la qualité s'explique par une importante corrélation entre les différentes couches induisant des confusions dans la classification. En conséquence, il fut décidé de conserver les données texturales de dissimilarité, de moyenne GLCM et de la variance GLCM, ces trois données étant peu corrélées. De plus, des tests statistiques ont permis d'identifier la donnée texturale de dissimilarité comme la plus significative pour la discrimination des forêts ayant connu de récentes exploitations de celles ayant connues de plus anciennes exploitations.

La combinaison des données texturales avec les données SAR-C et SRTM permet l'obtention d'une classification à 62% d'exactitude sans toutefois être capable de discriminer les deux types forestiers précédemment cités. Ces résultats de faible qualité témoignent de l'incapacité des méthodes de traitements appliquées dans le cadre de cette étude, à réaliser une discrimination correcte des écosystèmes forestiers via la classification par pixel.

Les segmentations d'images, réalisées en vue d'une classification par objet, se sont déroulées de manière stratifiée ; les segmentations d'un seul tenant ne permettant par l'extraction des objets d'intérêts de manière qualitative. Des différents algorithmes de segmentations testés, l'algorithme MRS ainsi que son homonyme automatisant la recherche du paramètre d'échelle optimal MRS-ESP2 furent évalués comme étant les plus efficaces pour mettre en évidence différents objets d'intérêts. Ces algorithmes diffèrent cependant par leur temps de calcul, MRS-ESP2 prenant beaucoup plus de temps suite à sa procédure de calcul par itération. Les classifications réalisées par objets en utilisant le même jeu de données que celles réalisées par pixels ont permis d'obtenir un F-score général de 72%.

Ces différents résultats démontrent le potentiel des données Sentinel-1 pour la cartographie forestière, et ce, en particulier pour la détection des milieux à régime hydrique particulier. Les données SAR-C se sont en revanche montrées insuffisantes à la discrimination des forêts ayant connu de récentes exploitations. Cependant, la combinaison de ces données avec celles acquises par les autres missions Sentinel de l'ESA permettrait la discrimination détaillée des différents écosystèmes forestiers tropicaux, les données optiques facilitant la détection des peuplements de différentes essences, tandis que les données radar permettraient la mise en évidence des zones marécageuses et humides caractérisées par une topographie particulière. Tenant compte de ces considérations, de plus amples recherches sur le sujet permettraient le développement d'une méthode de discrimination fine des écosystèmes forestiers pouvant s'appliquer à l'échelle du bassin du Congo.

Chapitre 6 : Perspective

Les résultats présentés au chapitre 4 ont pu mettre en exergue plusieurs éléments limitant la qualité des classifications. Afin d'améliorer la qualité de ces dernières, plusieurs pistes peuvent être suivies.

6.1 Indice radar de végétation (RVI)

Corrélié à la présence d'eau dans le feuillage et à l'indice de surface foliaire, cet indice est une métrique bien définie permettant d'évaluer le couvert végétal. À l'instar de son homonyme optique NDVI, l'indice RVI utilise les intensités des rétrodiffusions linéaires des polarisations simples (VV, HH) et croisées (HV, VH) pour quantifier la présence de végétation sur la surface d'intérêt. Cet indice, basé sur l'équation ci-dessous, nécessite dès lors l'acquisition de données en quadruple polarisation (Kumar et al., 2013). Indisponible pour les données S1, la quadruple polarisation est désormais utilisable grâce au récent lancement de la RADARSAT Constellation Mission (RCM), le 12 juin 2019 par SpaceX (SpaceX, 2019).

$$RVI = \frac{8\sigma_{HV}}{\sigma_{HH} + \sigma_{VV} + 2\sigma_{HV}}$$

Le ratio de ces différentes polarisations génère un résultat compris entre zéro et un, un correspondant au maximum de la couverture végétale. Ce résultat peut toutefois être supérieur à 1 dans le cas de rétrodiffusions par double rebond. Bien que cet indice fût principalement utilisé pour qualifier l'état de croissances de cultures (Kumar, et al., 2013; Haldar et al., 2018), d'autres études l'ont repris pour dresser des masques forêt/non-forêt ainsi que pour améliorer les opérations de quantification de la biomasse forestière (Mohan et al., 2011).

6.2 Angle d'incidence

Comme précisé dans la section précédente, l'angle d'incidence d'acquisition de la donnée influe sur les propriétés radiométriques de l'image résultante. En conséquence, de nombreuses applications privilégieront un angle d'incidence ou une gamme d'angles d'incidence bien précis. La qualification des ressources forestières fait partie de ces applications : l'utilisation d'un faible angle d'incidence permet d'améliorer la sensibilité du capteur à la quantification de la biomasse alors qu'un angle d'incidence plus important permet d'améliorer la discrimination des types forestiers par interaction avec la structure de la canopée. Les satellites Sentinel-1 acquièrent leurs données selon un angle d'incidence compris entre 29,1° et 46°. Il est donc possible, moyennant le choix des bonnes orbites, de pouvoir acquérir des données sur une même zone d'étude à des angles d'incidence différents. Cependant, l'acquisition de telles données n'est pas encore réalisable uniquement à partir des satellites Sentinel-1, ces derniers couvrant le continent africain de manière limitée en comparaison avec l'Europe (Figure 55). Les images acquises avec des angles d'incidence différents permettraient d'améliorer les classifications de couverture du sol. Néanmoins, ces images étant prises depuis des orbites différentes, à des moments différents, les composites formés incluront une composante de variation temporelle (ASAR Science Advisory Group, 1998).

Sentinel-1 Constellation Observation Scenario: Mode - Polarisation - Observation Geometry

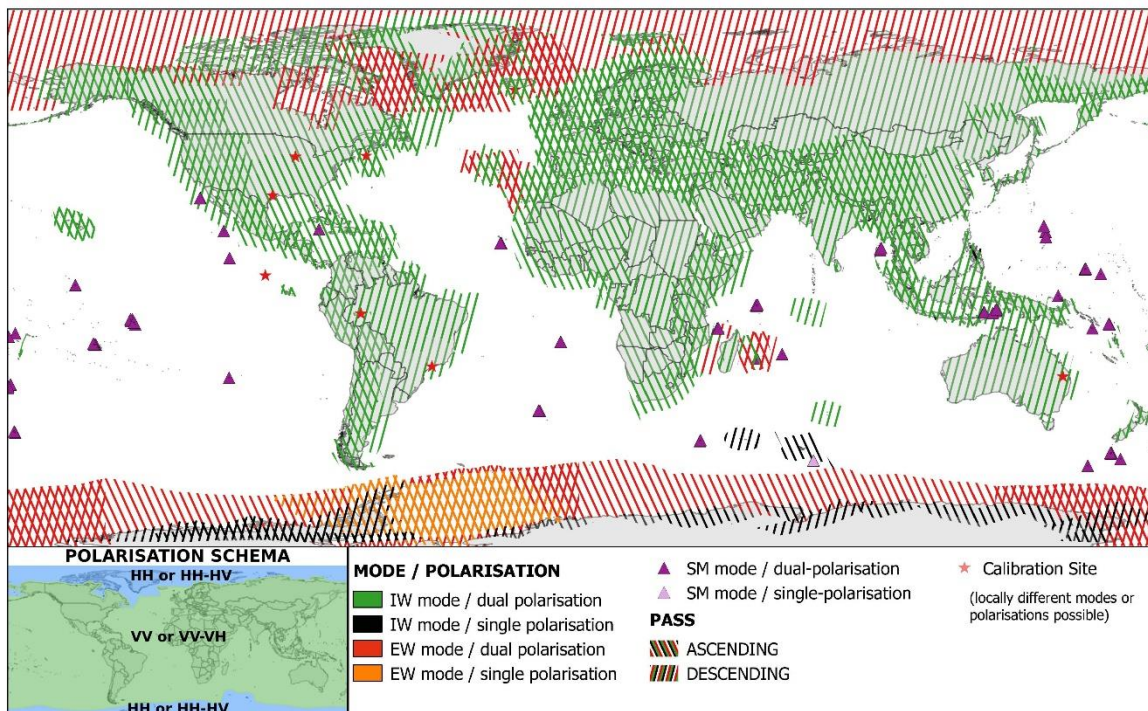


Figure 55 : Scénario d'observation des satellites Sentinel-1

6.3 Interférométrie et mesure de cohérence

L'interférométrie consiste en l'exploitation de plusieurs images radar prises à des périodes ou à des positions légèrement différentes pour en analyser les différences de phase. Utilisée principalement dans le domaine de la volcanologie, des glissements de terrains et tremblements de terre, l'interférométrie permet de détecter des changements à l'échelle millimétrique. En foresterie, l'interférométrie est régulièrement utilisée pour détecter la présence de coupes à blanc ou autres perturbations induisant des modifications de la structure de la canopée (Wulder et Franklin, 2006; G. Smith et Askne, 1997). Une étude menée par Treuhaft et al. (2017) suggère que l'utilisation de l'interférométrie permettrait la discrimination des forêts tropicales primaires et secondaires sur base d'une estimation de la biomasse aérienne. Une seconde étude décrit que l'analyse de la cohérence et de la phase permet une meilleure détermination des structures végétales forestières ainsi qu'une meilleure estimation de la biomasse aérienne que l'analyse des données radar sur l'intensité. Cette précision s'expliquerait par le fait que les signaux dus aux changements des paramètres de structures et de biomasse acquis par interférométrie seraient supérieurs au bruit des observations pour une plus large gamme de paramètres de structure et de biomasse que les observations réalisées à partir de l'intensité (R. N. Treuhaft et Siqueira, 2004).

Bibliographie

- Abdel-Hamid, A., Dubovyk, O., Abou El-Magd, I., Menz, G., 2018. Mapping Mangroves Extents on the Red Sea Coastline in Egypt using Polarimetric SAR and High Resolution Optical Remote Sensing Data. *Sustainability* 10, 646. <https://doi.org/10.3390/su10030646>
- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S., 2012. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Trans. Pattern Anal. Mach. Intell.* 34, 2274–2282. <https://doi.org/10.1109/TPAMI.2012.120>
- Agence spatiale européenne, n.d. Sentinel-2 - Overview - Sentinel Online [WWW Document]. URL <https://sentinel.esa.int/web/sentinel/missions/sentinel-2/overview> (accessed 10.19.18).
- Aisen, B., 2006. A Comparison of Multiclass SVM Methods [WWW Document]. URL <https://courses.media.mit.edu/2006fall/mas622j/Projects/aisen-project/index.html> (accessed 6.16.19).
- Alena, D., Milutin, M., M, H., Wagner, W., 2016. FOREST AREA DERIVATION FROM SENTINEL-1 DATA 8.
- Alsdorf, D.E., Smith, L.C., Melack, J.M., 2001. Amazon floodplain water level changes measured with interferometric SIR-C radar. *IEEE Trans. Geosci. Remote Sens.* 39, 423–431.
- ArcMap 10.3, 2016. Fonction Ombrage—Aide | ArcGIS for Desktop [WWW Document]. ArcMAP. URL <http://desktop.arcgis.com/fr/arcmap/10.3/manage-data/raster-and-images/hillshade-function.htm> (accessed 7.29.19).
- Arnfield, J.A., 2019. Koppen climate classification [WWW Document]. *Encycl. Br.* URL <https://www.britannica.com/science/Koppen-climate-classification> (accessed 5.29.19).
- ASAR Science Advisory Group, 1998. ASAR Science and Applications. R.A. Harris, European Space Agency.
- Baatz, M., Schäpe, A., n.d. Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation 12.
- Balzter, H., Cole, B., Thiel, C., Schmullius, C., 2015. Mapping CORINE Land Cover from Sentinel-1A SAR and SRTM Digital Elevation Model Data using Random Forests. *Remote Sens.* 7, 14876–14898. <https://doi.org/10.3390/rs71114876>
- Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Sci. Data* 5, 180214. <https://doi.org/10.1038/sdata.2018.214>
- Bergeri, I., Michel, R., Boutin, J.-P., 2002. POUR TOUT SAVOIR OU PRESQUE SUR LE COEFFICIENT KAPPA... 3.
- Bergonzini, J.-C., Lanly, J.-P., 2000. Les forêts tropicales. Editions Quae.
- Berthier, S., 2012. 3 - Polarisation effects in natural photonic structures and their applications, in: Large, M. (Ed.), *Optical Biomimetics*, Woodhead Publishing Series in Electronic and Optical Materials. Woodhead Publishing, pp. 79–115e. <https://doi.org/10.1533/9780857097651.79>
- Besag, J., 1986. On the Statistical Analysis of Dirty Pictures. *J. R. Stat. Soc. Ser. B Methodol.* 48, 259–279. <https://doi.org/10.1111/j.2517-6161.1986.tb01412.x>
- Blanc, L., Gond, V., Ho Tong Minh, D., 2016. 2 - Remote Sensing and Measuring Deforestation, in: Baghdadi, N., Zribi, M. (Eds.), *Land Surface Remote Sensing*. Elsevier, pp. 27–53. <https://doi.org/10.1016/B978-1-78548-105-5.50002-5>

- Bogdanov, A.V., Sandven, S., Johannessen, O.M., Alexandrov, V.Y., Bobylev, L.P., 2005. Multisensor approach to automated classification of sea ice image data. *IEEE Trans. Geosci. Remote Sens.* 43, 1648–1664. <https://doi.org/10.1109/TGRS.2005.846882>
- Breiman, L., 2001. Random Forest.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Image Classification using Statistical Learning Methods [WWW Document]. URL <https://rafalab.github.io/pages/649/section-11.pdf> (accessed 1.19.19).
- Brown, M., 1999. Fisher’s linear discriminant [WWW Document]. URL <https://compbio.soe.ucsc.edu/genex/genexTR2html/node12.html> (accessed 6.16.19).
- Bwangoy, J.-R.B., Hansen, M.C., Roy, D.P., Grandi, G.D., Justice, C.O., 2010. Wetland mapping in the Congo Basin using optical and radar remotely sensed data and derived topographical indices. *Remote Sens. Environ.* 114, 73–86. <https://doi.org/10.1016/j.rse.2009.08.004>
- Cao, Y., Sun, H., Xu, X., 2005. An Unsupervised Segmentation Method Based on MPM for SAR Images. *IEEE Geosci. Remote Sens. Lett.* 2, 55–58. <https://doi.org/10.1109/LGRS.2004.839649>
- Cassol, H.L.G., Shimabukuro, Y.E., Beuchle, R., 2019. SENTINEL-1 TIME-SERIES ANALYSIS FOR DETECTION OF FOREST DEGRADATION BY SELECTIVE LOGGING 5.
- Chambers, J.Q., Asner, G.P., Morton, D.C., Anderson, L.O., Saatchi, S.S., Espírito-Santo, F.D.B., Palace, M., Souza, C., 2007. Regional ecosystem structure and function: ecological insights from remote sensing of tropical forests. *Trends Ecol. Evol.* 22, 414–423. <https://doi.org/10.1016/j.tree.2007.05.001>
- Chevassus-Au-Louis, B., Pirard, R., 2011. Les services écosystémiques des forêts et leur rémunération éventuelle. *Rev. For. Fr.* <https://doi.org/10.4267/2042/46106>
- Choi, H., Baraniuk, R.G., 2001. Multiscale image segmentation using wavelet-domain hidden Markov models. *IEEE Trans. Image Process.* 10, 1309–1321. <https://doi.org/10.1109/83.941855>
- Clauss, K., Ottinger, M., Kuenzer, C., 2018. Mapping rice areas with Sentinel-1 time series and superpixel segmentation. *Int. J. Remote Sens.* 39, 1399–1420. <https://doi.org/10.1080/01431161.2017.1404162>
- Clerici, N., Calderón, C.A.V., Posada, J.M., 2017. Fusion of Sentinel-1A and Sentinel-2A data for land cover mapping: a case study in the lower Magdalena region, Colombia. *J. Maps* 13, 718–726. <https://doi.org/10.1080/17445647.2017.1372316>
- CNTRL, 2012. TÉLÉ DÉTECTION : Définition de TÉLÉ DÉTECTION [WWW Document]. Cent. Natl. Ressour. Textuelles Lexicales. URL <https://www.cnrtl.fr/definition/t%C3%A9l%C3%A9d%C3%A9tection> (accessed 8.12.19).
- Cook, R., McConnell, I., Stewart, D., 1996. Segmentation and simulated annealing.
- Delgado-Aguilar, M.J., Fassnacht, F.E., Peralvo, M., Gross, C.P., Schmitt, C.B., 2017. Potential of TerraSAR-X and Sentinel 1 imagery to map deforested areas and derive degradation status in complex rain forests of Ecuador. *Int. For. Rev.* 19, 102–118. <https://doi.org/10.1505/146554817820888636>
- Drăguț, L., Csillik, O., Eisank, C., Tiede, D., 2014. Automated parameterisation for multi-scale image segmentation on multiple layers. *Isprs J. Photogramm. Remote Sens.* 88, 119–127. <https://doi.org/10.1016/j.isprsjprs.2013.11.018>

- Drăguț, L., Eisank, C., 2012. Automated object-based classification of topography from SRTM data. *Geomorphology* 141–142, 21–33.
<https://doi.org/10.1016/j.geomorph.2011.12.001>
- Drăguț, L., Tiede, D., Levick, S.R., 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *Int. J. Geogr. Inf. Sci.* 24, 859–871. <https://doi.org/10.1080/13658810903174803>
- Efron, B., 1982. The Jackknife, the Bootstrap and Other Resampling Plans.
- EPFL, 2018. SLIC Superpixels. URL <https://ivrl.epfl.ch/research-2/research-current/research-superpixels/> (accessed 6.3.19).
- ESA, 2019a. ALOS - eoPortal Directory - Satellite Missions [WWW Document]. Eo Portal Dir. URL <https://directory.eoportal.org/web/eoportal/satellite-missions/a/alos> (accessed 6.30.19).
- ESA, 2019b. Single Look Complex - Sentinel-1 SAR Technical Guide - Sentinel Online [WWW Document]. URL <https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-1-sar/products-algorithms/level-1-algorithms/single-look-complex> (accessed 6.2.19).
- ESA, 2018. Sentinel-1 new product format operational on 13 March 2018 [WWW Document]. News- Copernic. Open Access Hub. URL <https://scihub.copernicus.eu/news/News00301> (accessed 8.10.19).
- ESA STEP Forum, 2018. GRD-Border-Noise and Thermal noise removal are not working anymore since March 13. 2018 [WWW Document]. ESA STEP Forum. URL <https://forum.step.esa.int/t/grd-border-noise-and-thermal-noise-removal-are-not-working-anymore-since-march-13-2018/9332> (accessed 8.10.19).
- Fisher, P., 1997. The pixel: A snare and a delusion. *Int. J. Remote Sens.* 18, 679–685.
<https://doi.org/10.1080/014311697219015>
- Fosgate, C.H., Krim, H., Irving, W.W., Karl, W.C., Willsky, A.S., 1997. Multiscale segmentation and anomaly enhancement of SAR imagery. *IEEE Trans. Image Process.* 6, 7–20.
<https://doi.org/10.1109/83.552077>
- FSC, s.d. Vision et mission [WWW Document]. For. Steward. Counc. URL <http://www.fsc.be/fr-be/a-propos-du-fsc/vision-et-mission> (accessed 8.12.19).
- Gao, G., Zhao, L., Zhang, J., Zhou, D., Huang, J., 2008. A segmentation algorithm for SAR images based on the anisotropic heat diffusion equation. *Pattern Recognit.* 41, 3035–3043. <https://doi.org/10.1016/j.patcog.2008.01.029>
- Ghazaryan, G., Dubovyk, O., Löw, F., Lavreniuk, M., Kolotii, A., Schellberg, J., Kussul, N., 2018. A rule-based approach for crop identification using multi-temporal and multi-sensor phenological metrics. *Eur. J. Remote Sens.* 51, 511–524.
<https://doi.org/10.1080/22797254.2018.1455540>
- Gidudu, A., Hulley, G., Marwala, T., 2014. Image Classification Using SVMs: One-against-One Vs One-against-All 6.
- Gilmore, S., Saleem, A., Dewan, A., 2015. Effectiveness of DOS (Dark-Object Subtraction) method and water index techniques to map wetlands in a rapidly urbanising megacity with Landsat 8 data. *Res. Locate15* 100–108.
- Gond, V., Betbeder, J., Fayolle, A., Viennois, G., 2015. NEW INSIGHTS IN TROPICAL FOREST DIVERSITY MAPPING IN CENTRAL AFRICA USING LOW RESOLUTION REMOTE SENSING.
- Goodchild, M.F., 2007. Towards a general theory of geographic representation in GIS: *International Journal of Geographical Information Science: Vol 21, No 3* [WWW Document]. URL

- <https://www.tandfonline.com/doi/full/10.1080/13658810600965271> (accessed 6.11.19).
- Goodman, J.W., 2007. *Speckle Phenomena in Optics: Theory and Applications*. Roberts and Company Publishers.
- Grover, K., Quegan, S., da Costa Freitas, C., 1999. Quantitative estimation of tropical forest cover by SAR. *IEEE Trans. Geosci. Remote Sens.* 37, 479–490.
<https://doi.org/10.1109/36.739096>
- Haldar, D., Dave, R., Dave, V.A., 2018. Evaluation of full-polarimetric parameters for vegetation monitoring in rabi (winter) season. *Egypt. J. Remote Sens. Space Sci., EJRS Special Issue: Microwave Remote Sensing in honor of Professor Adel Yehia* 21, S67–S73. <https://doi.org/10.1016/j.ejrs.2018.05.002>
- Hall-Beyer, M., 2017. GLCM Texture: A Tutorial v. 3.0 March 2017.
<http://dx.doi.org/10.11575/PRISM/33280>
- Haralick, R., n.d. Textural Features for Image Classification [WWW Document]. URL <http://haralick.org/journals/TexturalFeatures.pdf> (accessed 4.9.19).
- Hastie, T., Tibsharani, R., Friedman, J., 2009. *The Elements of Statistical Learning*. JAXA, 1997. About ALOS [WWW Document]. URL https://www.eorc.jaxa.jp/ALOS/en/about/about_index.htm (accessed 5.7.19).
- Jennings, S.B., Brown, N.D., Sheil, D., 1999. Assessing forest canopies and understorey illumination: canopy closure, canopy cover and other measures. *For. Int. J. For. Res.* 72, 59–74. <https://doi.org/10.1093/forestry/72.1.59>
- Kumar, S.D., Rao, S.S., Sharma, J.R., 2013. Radar Vegetation Index as an Alternative to NDVI for Monitoring of Soyabean and Cotton 7.
- Lardeux, C., Frison, P.-L., Rudant, J.-P., Souyris, J.-C., Tison, C., Stoll, B., 2006. Use of the SVM Classification with Polarimetric SAR Data for Land Use Cartography, in: 2006 IEEE International Symposium on Geoscience and Remote Sensing. Presented at the 2006 IEEE International Symposium on Geoscience and Remote Sensing, IEEE, Denver, CO, USA, pp. 493–496. <https://doi.org/10.1109/IGARSS.2006.131>
- Le Toan, T., Picard, G., Martinez, J.-M., Melon, P., Davidson, M., 2002. On the relationships between radar measurements and forest structure and biomass. Presented at the Retrieval of Bio- and Geo-Physical Parameters from SAR Data for Land Applications, pp. 3–12.
- Lewis M.Cowardin, 1979. *Classification of~ Wetlands _and~ Deepwater Habitats~ of the United States*.
- Liu, Y., Gong, W., Hu, X., Gong, J., 2018. Forest Type Identification with Random Forest Using Sentinel-1A, Sentinel-2A, Multi-Temporal Landsat-8 and DEM Data. *Remote Sens.* 10, 946. <https://doi.org/10.3390/rs10060946>
- Lohberger, S., Stängel, M., Atwood, E.C., Siegert, F., 2018. Spatial evaluation of Indonesia’s 2015 fire-affected area and estimated carbon emissions using Sentinel-1. *Glob. Change Biol.* 24, 644–654. <https://doi.org/10.1111/gcb.13841>
- Lombardo, P., Macri Pellizzeri, T., 2003. A Comparison of Statistical Segmentation Techniques for Polarimetric Sar: Region Growing Versus Simulated Annealing.
- MacMillan, R.A., Pettapiece, W.W., Nolan, S.C., Goddard, T.W., 2000. A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. *Fuzzy Sets Syst.* 113, 81–109.
[https://doi.org/10.1016/S0165-0114\(99\)00014-7](https://doi.org/10.1016/S0165-0114(99)00014-7)
- Mahafza, B.R., 2000. Chapter 2: Radar Cross Section (RCS) 46.

- Malik, C., Malhotra, R., 2016. A Survey on Various Classification Methods for SAR Images.
- Mayaux, P., Richards, T., Janodet, E., 1999. A vegetation map of Central Africa derived from satellite imagery. *J. Biogeogr.* 26, 353–366. <https://doi.org/10.1046/j.1365-2699.1999.00270.x>
- Mertens, B., Minnemeyer, S., Ayenika Nsoyuni, L., Steil, M., 2007. Rapport WRI : Atlas forestier interactif du Congo, Document de Synthèse, version 1.0.
- Mohan, S., Das, A., Haldar, D., Maity, S., 2011. Monitoring and retrieval of vegetation parameter using multi-frequency polarimetric SAR data, in: 2011 3rd International Asia-Pacific Conference on Synthetic Aperture Radar (APSAR). Presented at the 2011 3rd International Asia-Pacific Conference on Synthetic Aperture Radar (APSAR), pp. 1–4.
- Numbisi, F.N., Van Coillie, F., De Wulf, R., 2018. MULTI-DATE SENTINEL1 SAR IMAGE TEXTURES DISCRIMINATE PERENNIAL AGROFORESTS IN A TROPICAL FOREST-SAVANNAH TRANSITION LANDSCAPE. *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XLII–1, 339–346. <https://doi.org/10.5194/isprs-archives-XLII-1-339-2018>
- OFAC, 2017. OFAC : Synthèse nationale de la République du Congo [WWW Document]. *Obs. For. Afr. Cent.* URL https://www.observatoire-comifac.net/monitoring_system/national_synthesis/COG (accessed 8.2.19).
- Oliver, C.J., 2000. Rain forest classification based on SAR texture. *IEEE Trans. Geosci. Remote Sens.* 38, 1095–1104. <https://doi.org/10.1109/36.841988>
- Onkagui, J., n.d. La gestion durable des forêts en République du Congo [WWW Document]. URL <http://www.fao.org/docrep/ARTICLE/WFC/XII/1024-B4.HTM> (accessed 10.17.18).
- Ouaidrari, H., Vermote, E.F., 1999. Operational atmospheric correction of Landsat TM data. *Remote Sens. Environ.* 70, 4–15.
- Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. *Hydrol Earth Syst Sci* 12.
- PFBC, 2012. Congolaise Industrielle des Bois (CIB) / Groupe OLAM international.
- Piantanida, R., 2017. Thermal Denoising of Products Generated by the S-1 IPF 25.
- Rüetschi, M., Schaepman, M.E., Small, D., 2018. Using Multitemporal Sentinel-1 C-band Backscatter to Monitor Phenology and Classify Deciduous and Coniferous Forests in Northern Switzerland. *Remote Sens.* 10, 55. <https://doi.org/10.3390/rs10010055>
- Ryherd, S., Woodcock, C., 1996. Combining Spectral and Texture Data in the Segmentation of Remotely Sensed Images 14.
- Sakarya, U., Demirpolat, C., 2018. SAR image time-series analysis framework using morphological operators and global and local information-based linear discriminant analysis 9.
- Shastri, R.K., 2015. A STUDY OF SPECKLE NOISE REDUCTION FILTERS.
- Shell, J.R., 2004. Bidirectional Reflectance : An Overview with Remote Sensing Applications & Measurement Recommendations 80.
- Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R., Lucas, R., 2014. New global forest/non-forest maps from ALOS PALSAR data (2007–2010). *Remote Sens. Environ.* 155, 13–31. <https://doi.org/10.1016/j.rse.2014.04.014>
- Small, D., 2011. Flattening Gamma: Radiometric Terrain Correction for SAR Imagery. *IEEE Trans. Geosci. Remote Sens.* 49, 3081–3093. <https://doi.org/10.1109/TGRS.2011.2120616>

- Smith, G., Askne, J., 1997. Use of ERS SAR interferometry for monitoring clear-cutting of forests, in: IGARSS'97. 1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing - A Scientific Vision for Sustainable Development. Presented at the IGARSS'97. 1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing - A Scientific Vision for Sustainable Development, pp. 793–796 vol.2.
<https://doi.org/10.1109/IGARSS.1997.615259>
- Smith, G., Beare, M., Boyd, M., Downs, T., Gregory, M., Morton, D., Brown, N., Thomson, A., 2007. UK Land Cover Map Production Through the Generalisation of OS MasterMap®. *Cartogr. J.* 44, 276–283. <https://doi.org/10.1179/000870407X241827>
- SpaceX, 2019. RADARSAT CONSTELLATION MISSION.
- Treuhaft, R., Lei, Y., Gonçalves, F., Keller, M., Santos, J., Neumann, M., Almeida, A., 2017. Tropical-Forest Structure and Biomass Dynamics from TanDEM-X Radar Interferometry. *Forests* 8, 277. <https://doi.org/10.3390/f8080277>
- Treuhaft, R.N., Siqueira, P.R., 2004. The calculated performance of forest structure and biomass estimates from interferometric radar. *Waves Random Media* 14, S345–S358. <https://doi.org/10.1088/0959-7174/14/2/013>
- van der Sanden, J.J., Hoekman, D.H., 1999. Potential of Airborne Radar To Support the Assessment of Land Cover in a Tropical Rain Forest Environment. *Remote Sens. Environ.* 68, 26–40. [https://doi.org/10.1016/S0034-4257\(98\)00099-6](https://doi.org/10.1016/S0034-4257(98)00099-6)
- Vapnik, V., 1995. *The Nature of Statistical Learning Theory*. Springer-Verlag, London.
- Verhegghen, A., Eva, H., Desclée, B., Achard, F., 2016. Review and combination of recent remote sensing based products for forest cover change assessments in Cameroon. *Int. For. Rev.* 18, 14–25. <https://doi.org/10.1505/146554816819683807>
- Wang, Y., Hess, L.L., Filoso, S., Melack, J.M., 1995. Understanding the radar backscattering from flooded and nonflooded Amazonian forests: Results from canopy backscatter modeling. *Remote Sens. Environ.* 54, 324–332.
- Wood, J., 1996. *The geomorphological characterisation of Digital Elevation Models*. ProQuest.
- Wulder, M.A., Franklin, S.E., 2006. *Understanding Forest Disturbance and Spatial Pattern: Remote Sensing and GIS Approaches*. CRC Press.
- Xu, H., Wang, W., Liu, X., 2011. A fast segmentation approach of SAR image by fusing optical image, in: 2011 IEEE International Geoscience and Remote Sensing Symposium. Presented at the 2011 IEEE International Geoscience and Remote Sensing Symposium, pp. 2665–2668. <https://doi.org/10.1109/IGARSS.2011.6049751>
- Zhao, K., Suarez, J.C., Garcia, M., Hu, T., Wang, C., Londo, A., 2018. Utility of multitemporal lidar for forest and carbon monitoring: Tree growth, biomass dynamics, and carbon flux. *Remote Sens. Environ.* 204, 883–897. <https://doi.org/10.1016/j.rse.2017.09.007>

Développement d'une méthode de traitement des séries temporelles des satellites Sentinel-1 permettant la discrimination des écosystèmes forestiers du bassin du Congo

Julien Herinckx

Comprise dans des latitudes inférieures à 30° nord et sud de l'équateur, les écosystèmes forestiers tropicaux couvrent 13,5 % des terres émergées et abritent l'une des biodiversités les plus riches sur terre. Bien que ces milieux soient régulièrement étudiés, l'état des connaissances des forêts tropicales reste lacunaire. Ces lacunes s'expliquent principalement par la difficulté d'acquisition de données sur ces milieux, les forêts tropicales étant des écosystèmes complexes à l'accessibilité limitée.

Les techniques d'acquisition de données par télédétection permettent de répondre à cette contrainte d'accessibilité. Le récent lancement des satellites européens Sentinel-1 et Sentinel-2 permet la collecte de données synoptiques, systématiques et avec une grande répétition temporelle offrant ainsi une excellente opportunité pour travailler sur ces milieux de vaste étendue.

Ce mémoire se concentre sur l'exploitation des données des satellites Sentinel-1 en vue de produire une discrimination des écosystèmes forestiers du bassin du Congo. Afin de réaliser une telle discrimination, deux méthodes de classification sont testées.

La première, par pixel, résulte en une classification aux performances modérées, ne suffisant pas à réaliser une discrimination correcte des différents écosystèmes forestiers.

La seconde, par objet, se base sur l'utilisation de deux algorithmes de segmentation : l'algorithme multirésolution et multirésolution avec estimation du paramètre d'échelle. La classification des objets extraits de la segmentation est, quant à elle, réalisée via l'algorithme Random Forest et génère une carte présentant de bonnes performances (F-score : 72%), en particulier pour l'identification des milieux marécageux.

Les différents résultats obtenus lors de cette étude témoignent de la bonne capacité des données Sentinel-1 à réaliser la discrimination des écosystèmes forestiers tropicaux.